

TEACHER SELECTION: USE OF DEMONSTRATION LESSONS

A Dissertation  
Presented to  
the School of Education  
Drake University

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Education


by Richard J. Wede  
June 1996


TEACHER SELECTION: USE OF DEMONSTRATION LESSONS

by Richard J. Wede

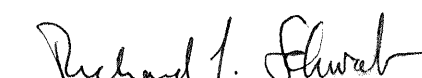
June 1996

Approved by Committee:

  
Dr. Annette M. Liggett, Chair

  
Dr. Richard L. Schwab

  
Dr. James Cox

  
Dr. Richard L. Schwab  
Dean of the School of Education

1996  
.W414

ACG 9014

## TEACHER SELECTION: USE OF DEMONSTRATION LESSONS

An abstract of a Dissertation by  
Richard J. Wede  
June 1996  
Drake University  
Advisor: Annette M. Liggett

The problem. This study examined the use of a teacher demonstration lesson (TDL) as a component of one district's teacher selection process by determining if differences existed between evaluation ratings, as measured by the Summative Evaluation Report (SER), for teachers who were selected with the TDL and those who were selected without the TDL.

Procedure. The evaluation ratings of 101 teachers hired in a Midwest district were analyzed. The SER was divided into four performance areas and a total composite score. A Chi square test was used to identify significant differences for each area and the composite score between the 53 teachers hired from 1980 to 1983 and the 48 teachers hired from 1985 to 1988.

Findings. The research found no significant values for the four categories on the Summative Evaluation Report (SER) or for the composite score comparing the teachers selected using the TDL component and those selected without using the TDL. Therefore, the data failed to reject the null hypothesis.

Conclusions. The research identified that evaluation ratings of tenured staff did not discriminate sufficiently to identify differences among staff. The results of the study did not indicate that better qualified teachers were identified through the use of a TDL within the selection process. A review of the literature indicated that the selection process continues to be based primarily upon subjective criteria rather than objective criteria and that teacher performance evaluations might be improved through the use of authentic assessment methods.

Recommendations. Future studies regarding teacher selection procedures involving the TDL should investigate the evaluations for probationary teachers and the number who continue to be employed as tenured staff. In addition, a qualitative study regarding the perceptions of those involved in the selection process should be completed.

## ACKNOWLEDGMENTS

The completion of this paper was accomplished through the assistance of many significant people. The numerous valuable comments and suggestions that Dr. Richard Schwab and Dr. James Cox provided as members of my committee were greatly appreciated. Dr. Robert Hoehle's help in developing the study's design was invaluable. As my major advisor, Dr. Annette Liggett provided the important and necessary guidance and support that allowed me to accomplish the goal. To my good friend, Dorla Hill, for her comments and editing skills, thank you. To these individuals I owe a huge personal debt.

The technical advice and assistance by those individuals were extremely important in completing this paper; however, it was the support and sacrifices of my family that permitted me the time and solitude to complete it. A special thank you to my wife, Carol, who knew all along that I would finish even when I was not sure. I am especially grateful to my parents for the values taught and modeled and the continual encouragement to set high goals and work to achieve them.

## Table of Contents

	Page
List of Table. . . . .	vi
List of Figure . . . . .	vi
Chapter	
1. Introduction . . . . .	1
Purpose of the Study . . . . .	1
Problem Statement . . . . .	2
Framework of the Study . . . . .	3
Assumptions . . . . .	4
Limitations . . . . .	4
Delimitation. . . . .	5
Definition of Terms. . . . .	6
Organization of the Study. . . . .	7
2. Review of the Literature. . . . .	8
Introduction and Scope. . . . .	8

	Teacher Influence on Student Learning . . . . .	8
	Evaluation Instruments . . . . .	11
	Hiring Practices . . . . .	23
	Summary. . . . .	45
3.	Design of the Study . . . . .	47
	Independent Variables . . . . .	48
	Dependent Variable. . . . .	51
	Population . . . . .	53
	Rights of Human Subjects. . . . .	54
	Instrumentation . . . . .	54
	Data Collection . . . . .	56
	Treatment of the Data. . . . .	56
	Null Hypothesis . . . . .	57
	Summary . . . . .	57
4.	Analysis of the Data . . . . .	59
	Item Analysis and Results. . . . .	61
	Summary . . . . .	63
5.	Summary, Findings, Implications, and Recommendations. . . .	64
	Summary . . . . .	64
	Findings. . . . .	65

Implications. . . . .	65
Recommendations. . . . .	69
References . . . . .	71
Appendices	
A. Summative Evaluation Report for School District A . . . . .	81

## Table

		Page
Table		
1.	Chi-Square Test of Observed and Expected Evaluation Frequency Counts for Non-TDL and TDL Teacher Groups . . . . .	62

## Figure

		Page
Figure		
1.	Counts for Non-TDL and TDL Teacher Groups . . . . .	48



## Chapter 1

### INTRODUCTION

Vann (1989), Herman and Stephens (1987), Caliendo (1986), Cardozier (1985), and others have suggested that the quality of schools is primarily dependent upon the quality of teachers employed by the district. Teacher selection is one of the most important decisions that a school administrator can make (Vann, 1989; Wise, Darling-Hammond, & Berry, 1987; Moore, 1987; Smith, 1980). Past selection practices that involved only paper screening and personal interviews oftentimes limited administrators in identifying the most appropriate faculty members for their schools.

Credentials and interviews are only a portion of the selection process that administrators can use to learn what the applicant and others say that person can do. An effective selection process should provide district personnel with opportunities to hire the most qualified teachers. If districts are serious about improving student learning and promoting excellence within the classroom, then it may be appropriate to require applicants to actually demonstrate their teaching abilities in addition to using the perceptions of others and what the applicant says in an interview. Requiring prospective teachers to demonstrate their knowledge and skills through the presentation of a teaching lesson may improve the likelihood that administrators would hire the most appropriate candidates.

#### Purpose of the Study

Because of the importance of actually seeing how teachers teach before

they are hired, one would expect to find numerous studies concerning this topic. Browne (1972) and Stone (1972) conducted studies concerning the effect of observing candidates as they demonstrated their teaching skills during the selection process, and recently LeTendre (1989) studied the perceptions of district administrators who conducted an observation of teaching ability during the interview. This limited research indicated a need to further study teacher selection procedures in an effort to discover if there are more effective methods of hiring the best individuals from the pool of candidates being considered by the district. A study of teacher selection procedures could provide insight to administrators who seek to identify candidates who closely align with district performance standards for teaching, and thus, presumably would enhance student learning over time.

The practical significance of this study was the possibility of developing a teacher selection process that would more adequately screen candidates so administrators could better determine during the selection process which individuals, according to district standards, were the best qualified teachers. Such practice could allow building administrators to spend their time more effectively and efficiently by assisting new teachers improve their instructional skills rather than having to remediate unskilled, poorly trained teachers, or removing less-than-satisfactory ones.

#### Problem Statement

This study centered on the importance of a teacher candidate's demonstrated teaching ability as a component of a school district's employment

process. It examined the use of a teacher demonstration lesson (TDL) by determining if there were differences in the summative evaluations of teachers hired who used the TDL and those hired without it.

### Framework of the Study

The study included 101 classroom teachers employed in a school district. The participants were selected from the district's seniority lists which identified their date of employment and current assignment. Data were collected through the school district's seniority lists, personnel files, and data base of evaluation ratings. Only those individuals hired during two time periods (1980-83 or 1985-88) who were classroom teachers between 1990 and 1993 were included in the study. Those who no longer worked in the district or those who had moved into a position, such as guidance counselors, consultants or administrators, were excluded from the study.

The study also investigated three years of evaluation ratings of teachers in a school district using its Summative Evaluation Report (SER). The district changed its selection process to include the TDL. During that same time period, it also modified its teacher evaluation instrument. For that reason, the evaluations of the participants during their probationary period could not be used. Evaluations of those teachers hired prior to those changes were compared with the evaluations of those teachers hired using the TDL. The data collection of those evaluation ratings were over a three year time span because the district's evaluation procedures required tenured staff to be evaluated only once every three years.

The research question for this study was:

Does the distribution of ratings on the SER for teachers employed using the TDL differ from those employed without using the TDL?

A Chi square test was used to analyze the nominal data collected.

#### Assumptions

Following are the assumptions that guided this study.

1. One of the most important administrative tasks for determining student learning was the selection of the best teachers.
2. The Summative Evaluation Report (SER) accurately reflected the district's expectations for teaching performance.
3. The Teacher Demonstration Lesson (TDL) accurately reflected the teaching ability of a teacher candidate and provided interviewers the opportunity to judge the lesson taught in relation to the school district's expectations for teachers.

#### Limitations

This study had three primary limitations.

1. Contexts of districts vary.
2. The participants in the study were limited due to the number of employment openings available in the district during the time of the study.
3. The educational level and professional experiences for applicants were different.

The first limitation involved the context of a particular district. The study focused on only one school district, comparing the hiring procedures of a school district prior to its incorporating a TDL as a part of the total selection process to the procedures after including the TDL. The results of this study were limited to the unique characteristics of the district being studied. Therefore, readers need

to be cautious in applying the results of this study to other school districts' selection processes.

The second limitation concerned the number of teaching openings available within the district during the time of the study. The number of participants available for the study was limited to those hired for the openings during those years and who remained as classroom teachers.

The third limitation to the study was the individual professional differences that the candidates possessed. There were considerable differences among the candidates' education and experience. For example, some individuals had recently graduated with a bachelor degree in education and only student teaching experience or had one or more years of experience. Others had more experience and graduate hours in education in specific fields, and some applicants had completed an advanced degree. Those differences in the applicants could have affected the administrators' initial selection as they had identified specific characteristics that they sought for a particular position. Individuals who had recently graduated or had taken additional education courses would have a knowledge base and possible practice incorporating newer teaching practices and student learning theories. If a small group of the individuals had specific training in a particular program that the school used, it may have influenced that initial screening.

#### Delimitation

In addition to the limitations of the study, there existed one delimitation concerning the research. That delimitation was that the study analyzed the

teacher evaluations of those employed between seven and twelve years after their initial employment. Because the school district modified its evaluation instrument between 1983 and 1985, it was necessary to use the most recent evaluations of those groups hired rather than the first two years of evaluations during the participants' probationary period. Since the district's evaluation procedures required administrators to evaluate tenured teachers only once every three years, the evaluation data needed to be collected over a three-year time span.

### Definition of Terms

For the purpose of this study, certain terms that are used required clarification for the reader. Therefore, the following definitions are offered.

Building-level Administrators - individuals who were principals or assistant principals during the time period of the study and who evaluated all teachers included in this study.

Educational Level - the highest college degree attained by the teachers plus the graduate hours, if any, earned beyond that degree.

School Improvement Model (SIM) - a project under the direction of Dr. Richard Manatt, Iowa State University, designed to improve teacher performance and administrative evaluation skills with the ultimate goal of improved student learning.

Selection Process - the steps the district used when considering individuals for possible employment. Those steps included the review of applicant-completed materials, telephone reference checks, the interview, and in some instances, the teacher demonstration lesson (TDL).

Summative Evaluation Report (SER) - the formal written document that building-level administrators used to assess a teacher's job performance (see Appendix A).

Teacher Demonstration Lesson (TDL) - a teacher candidate presenting a 20-30 minute teaching vignette demonstrating his/her teaching ability. The TDL is taught under simplified conditions and is limited to the planning and presentation of a lesson. The lesson is presented to a small group of students.

Teachers - experienced and inexperienced individuals who completed the application process, possessed the appropriate teaching endorsement/ approval area, and were employed by the school district as licensed employees.

### Organization of the Study

Chapter I provides an introduction, indicating a purpose and need for the study. Chapter II discusses teaching as it relates to student learning, the evaluation process as it relates to the district's expectation of teachers, and the selection practices used by school personnel. Chapter III describes the methodology of the study, detailing the process used to collect and analyze the data. Chapter IV presents the data and the statistical analysis of the results. Chapter V provides a summary of the study's results, the conclusions reached, and recommendations for the future.

## Chapter 2

### REVIEW OF THE LITERATURE

#### Introduction and Scope

The purpose of this review was to address the following three questions that guided the framework of this study:

1. To what degree do teachers influence whether students will achieve success or failure in the classroom?
2. Can district administrators determine whether appropriate hiring decisions were made by studying the results of their teacher evaluation procedures?
3. Are current hiring practices effectively identifying those candidates who possess the qualities desired by the district?

#### Teacher Influence on Student Learning

The quality of schools is dependent upon classroom teaching; thus, an administrator who hires teachers has a significant impact on student learning by selecting excellent teachers (Herman & Stephens, 1987; Boyles & Engel, 1986; Bredeson, 1985; Cardozier, 1985). The teachers a school district employs essentially determine student learning (Carrick, 1980); yet, Tractenberg (1973) found that the broad consensus is a general discontent with the traditional way districts select or hire teachers. Several articles in recent years have concluded that many of today's teachers fail to demonstrate adequate skills and knowledge of the subject matter for the classroom (Raths, 1989; Travers, 1989; Wise, Darling-Hammond, & Berry, 1988). Administrators must make a strong professional commitment to the improvement of public education by implementing research-based procedures for the selection of teachers (Mickler



& Solomon, 1986). It is the teacher, not the instructional materials or textbook, that can motivate and encourage students to develop their innate abilities.

I've come to a frightening conclusion that I am the decisive element in the classroom. It's my personal approach that creates the climate. It's my daily mood that makes the weather. As a teacher, I possess a tremendous power to make a child's life miserable or joyous. I can be a tool of torture or an instrument of inspiration. I can humiliate or humor, hurt or heal. In all situations, it is my response that decides whether a crisis will be escalated or de-escalated and a child humanized or dehumanized (Ginott, 1972, pp. 15-16).

Improving the teacher selection process greatly increases the possibility of hiring good teachers who will motivate students to learn.

Recruiting and selecting teachers may be the most important task school administrators perform. The quality of any school district depends more upon the quality of its staff than upon any other factor. Each time a teacher is hired, the local school and its district have an opportunity to improve instructional programs (Jensen, 1987, p. 5).

If the mission of schools is to provide excellent educational opportunities, then one of the most important decisions for district personnel is to select quality teaching staffs (Vann, 1989; Cardozier, 1985; Gerwin, 1974). Al-Rubaiy (1993) stated that the goal of the hiring procedure should be to select the best applicant for the job. The "best" person for a position was that individual who had the greatest chance to maximize student learning (Title, 1995). Ross (1991) stated, "Schools need high-quality teachers in the classroom. No other element of the educational process is as crucial for students to succeed" (p. 19). "It's a truism that the quality of the teacher is the most important determinant of quality in the classroom" (Cox, 1981, p. 3). Individual teachers have a more

significant affect on student achievement than either curriculum or method alone. It is the enthusiasm and competency of the teacher that is the essential ingredient which promotes student learning. In 1991, Oklahoma City Schools' superintendent, Arthur Steller (Clark, 1992), closed an elementary school as a result of its being placed on the state's "at-risk" list due to poor student performance on the Iowa Test of Basic Skills (total student composite scores were at the 49th percentile). Susan Clark was hired as the new principal and given the latitude and responsibility of hiring a totally new staff, including all teachers and support personnel. At the end of one year, student composite scores were at the 75th percentile, an increase of 26 percentile points. Clark spent many hours working with staff in a positive and caring manner, which in turn was reflected by the total staff's attitude toward the students. They truly believed that all students could learn. She attributed their success to outstanding teamwork and a positive attitude. "Steller credits much of the improvement in the troubled schools ("about 80 percent," he says) to changes in staff and school climate" (Harrington-Lueker, 1992, p. 26).

Administrators must use a variety of teacher selection approaches that will more likely influence student learning and achievement in a positive manner (Ghysels, 1988). All too often, however, administrators' selection decisions have been influenced by such factors as first impressions, personal attributes, and "the good old boy network", which have very little relationship to teaching performance (Young & McMurry, 1986). "If educators employed research-based information more often in their decision-making, perhaps the

schools would be able to function more effectively and provide better learning experiences for students" (Mickler & Solomon, 1986, p. 402). Gilbert and Lang (1967) suggested that administrators use measures of classroom performance derived from observations as a data source for improving teacher selection. Improving the selection process can increase the effectiveness and productivity of the student learning environment (Webster, 1988; Orebaugh, 1974).

### Evaluation Instruments

A renewed focus on the American educational scene has created considerable interest in teacher evaluation. Davey (1991) stated that the movement toward accountability has caused "monitoring and evaluation of teacher competency" to "become a key point of focus" (p. 121). He contended that "developing effective systems for assessing teacher competence has proven extremely difficult" (p. 121). Haertel (1986) and Ayers and Qualls (1979) determined that standardized tests indicate only a moderately strong relationship to success in teacher education programs; however, there was minimal validity concerning the actual day-to-day performance as a teacher.

Borich and Fenton (1977) stated that, ultimately, teachers should be evaluated according to their achievement of specifics, not on the basis of their standing compared to others. They stressed the importance of the evaluation process matching itself to the desired teacher behaviors sought by the district. When that fails to occur, then the procedure should be revised. Through a review of the literature, Kelly (1989) identified four behavioral variables that influenced and promoted effective teaching: a strong sense of the educational

and instructional goals, student centeredness, use of a variety of teaching methodologies, and appropriate and evaluative feedback. In comparing 74% of the evaluation instruments used in Iowa schools during the 1987-88 school year, Kelly determined that items addressing the critical issues represented 56% of the total number of items. This was significantly above the 50% level; however, it was considered relatively weak. Kelly concluded that evaluation instruments could be used more effectively by keeping in mind the four variables of effective teaching. She also determined that no relationship existed between the frequency of the evaluation process and the four dependent behavioral variables. Millman and Darling-Hammond (1990) determined that an evaluation system should be developed to mirror the district's view of teaching. It should be an intense appraisal of the teacher and occur at least every three years.

Burry and Shaw (1988) stressed the importance of evaluator training. Without it, teacher evaluation was suspect, and could be an invalid procedure. Evertson and Burry (1989) reported on the importance of classroom observations and the need for consistent, reliable data collection. Millman and Darling-Hammond (1990) echoed those points in viewing the classroom observation as a focal point of the systematic evaluation program. They also recognized that it had limitations, however, viewing it as necessary for providing the required direct view of seeing the teacher in action. The common instruments used in classroom observations have been brief narratives, rating scales, and checklists. They identified the following problems when using these

items:

1. The response alternatives on ordinal scales are unequal.
2. The lower end of the rating scale denoted unacceptable behavior which contributed to the tendency of raters to use the upper end of the scale.
3. Rater bias.
4. Lack of specificity.

All of these items could lead to measurement error.

McGreal (1990) conducted a study on the rating scales of the evaluation system. He found little benefit for schools to use the information from such scales. The scales do not provide the necessary specificity or discriminatory power needed within the evaluation procedure. He recommended the development of effective written summations. These would provide the evaluator and evaluatee with a more detailed recap of what occurred in the classroom, which would allow for a focused approach to the identification of what is needed within the classroom. Borich and Fenton (1977) contended, however, that inter-rater reliability was adequate when using the Lickert scale on evaluations if adequate time and sufficient resources were provided in the training of evaluators. They indicated that the Lickert scale was moderate-to-high in reliability but concurred that its validity may vary. Haefele (1992) stated that teacher evaluation was one of the most important procedures for verifying the quality of the educational system. When teacher evaluation was ineffective, it promoted an incompetent teaching force and failed to promote effective staff development or recognize outstanding teaching performance. Scriven (1990)

stated that the best predictor of job performance was how an individual performed on job-related tasks, not on a combination of group characteristics. Effective evaluations must center on what the individual was doing, not on preconceived opinions. Murphy (1987) supported the concept that instructional supervision and evaluation was the key for improving the quality of teaching and learning within the school. To be effective, he contended that the evaluation should be conducted over a three-year period. This would provide the needed time to collect and analyze data. Through this procedure, one-fourth of the staff would be out of the cycle for a year.

Dale (1991) stated, "A good selection system should result in hiring people on the basis of their qualifications, not because of criteria unrelated to the job" (p. 21). He believed that school districts could improve education if the reliability of hiring decisions were tested over time. Present employee performance must be compared to the rating that an individual received during the initial employment interviews. If ratings were inconsistent, then a review of the hiring system must be completed.

Authentic assessment of teachers has been advocated by many who want to see evaluation become more closely related to the actual tasks that teachers perform each day in the classroom. This new wave was heartily endorsed by Wiggins (1989) who wrote:

I propose. . . a return to the roots; we have lost sight of the fact that a true test of intellectual ability requires the performance of exemplary tasks. First, authentic assessments replicate the challenges and standards of performance that typically face writers, businesspeople, scientists, community leaders, designers

or historians. These include writing essays and reports, conducting individual and group research, designing proposals and mock-ups, assembling portfolios, and so on (p. 703).

Davey (1991) stated that authentic assessment can be accomplished "through the use of simulations, situational questions, problems embedded in practice, display of portfolios, and actual performance of the work" (p. 122). A valuable characteristic of performance assessment was that it appeared to be more equitable toward minority groups. Davey (1991) and Schmitt, Gooding, Noe, and Kirsch (1984) conducted studies that yielded similar scores for different ethnic groups, rather than large gaps in scores on standardized multiple-choice tests.

A critical component of performance assessments was determining the content domain (Davey, 1991). A job analysis was necessary in developing the performance assessment tasks since the instrument was to simulate important job functions. It was this information that served as the blueprint for the development of the exercise and its scoring rubrics. Although many did not believe it was important to carefully define the characteristics of effective teaching, two major problems occurred if left undefined. First, effective performance could not be demonstrated by teachers if effective teaching had not been defined in operational terms. Second, assessors might disagree on the standards unless consensus was reached.

Grover (1991) identified five significant advantages for use of the assessment center as a method of evaluating teachers (p. 113).

1. Articulates performance standards and makes the criteria public.

2. Reflects changes in the function and purpose of assessment as a result of changes in the view of teaching and learning.
3. Broadens the knowledge base and perspectives of all who participated in the evaluation process.
4. Adjusts to the different stages incurred during the transition as the gap that exists narrows over time.
5. Facilitates the changes identified as necessary to produce reform.

She contended that "the nature of the assessment center forces the articulation of dimensions and standards for effective performance (p. 113). Performance assessments resulted in objective, "sophisticated professional judgements that can serve as diagnostic as well as selection purposes" (p. 113).

In 1982, Richard Manatt, Iowa State University, developed the School Improvement Model (SIM) project which provided a complete array of recommended teacher performance areas, criteria response modes, and standards. The project's purpose was to improve teacher and administrative performance with the ultimate goal of improved student learning. Evaluators were provided time to learn how to effectively observe and document what occurred in the classroom (Manatt, 1982). Selected districts developed teacher and administrator teams that worked with Manatt and other educators from Iowa State University to develop staff inservice programs for teachers and administrators. The programs focused on student performance criteria, teacher performance, and teacher assessment through administrative evaluations. Using identified criteria of effective teaching behaviors, individual districts developed new teacher evaluation instruments with the goal of improved



measurement of teacher performance. The assessment involved multiple appraisers, student feedback, formative and summative steps, modified clinical supervision cycles, and a written agreement between the teacher and evaluator for improved performance in the next appraisal cycle. Characteristic of the SIM were components that encouraged teachers to set goals and standards for student performance. Teachers incorporated the goals and standards in their teaching and as assessment of student performance. This assessment was not only used for grading and evaluating students, but also to describe individual learning difficulties and prescribe specific remediation and reteaching procedures. Administrators reviewed these components within the teacher performance evaluations.

The Teacher Assessment Project (TAP), under the direction of Lee Shulman at Stanford University, was a research and development project funded by the Carnegie Corporation (Vavrus & Collins, 1991). The project personnel conducted developmental studies and field tests to determine the feasibility of performance-based teacher assessment methods. They chose to study assessment center exercises and performance portfolios. Both items required the researchers to focus on actual teaching practices, a process often omitted from other research projects. Fifty teachers, thirty elementary and twenty high school, participated in the TAP portfolios study. The primary objective of the study was to attempt an on-site documentation with considerable variability in conditions. Although validity and reliability tests were not addressed, the researchers provided new insights on how the

knowledge and skills of teachers might be assessed. The TAP determined that portfolios offered two advantages for teacher assessment: 1) an opportunity to capture context and 2) to present change and growth. Although the portfolio exercises were costly and difficult to design, standardize, and evaluate; the project was a worthwhile endeavor, paving the way for the development of new assessment forms that authentically identified and discriminated among the individual teacher qualities.

The TAP's second area of study was teacher assessment. TAP's intent was to explore new modes of teacher assessment. It was expected that these experiences would more accurately reflect the divergence and complexity of the teaching process than the traditional methods. In 1986-87, the TAP personnel developed and tested exercises for elementary mathematics teachers and high school history teachers. The participants were expected to demonstrate and explain their knowledge and skill in hypothetical situations similar to actual practice. During the exercises, participants "analyzed textbooks, evaluated student work, and analyzed video tapes of each other teaching in light of their own experiences" (Vavrus & Collins, 1991, p. 17).

While classroom observations may fail to accurately evaluate many of the critical dimensions of teaching, the performance assessment process was expensive and time-consuming to use. As used in the Stanford project, it also isolated the participants from their schools, colleagues, and students--the basis for their actual contexts of teaching. Because of the importance that was placed on the teacher's elocutionary skills, teachers who were better able to verbally

explain their teaching had an advantage. Shulman's work did help to emphasize what had been known by teachers for some time: "good teaching is the effective combination of knowledge with an understanding not so much of what we traditionally call methods, but of those key elements of a subject area that make information hang together for learners" (Tomala & Weinland, 1991, p. 194). They believed that Shulman caused educators to re-examine the understanding of classroom teaching to distinguish between activities that engage students and academic contexts that teachers used to help students develop understanding from what they learn.

The Connecticut Teacher Assessment Center (CONNTAC) Program attempted "to develop an integrative performance assessment process that will measure the knowledge skills, and instructional repertoire of beginning teachers" (Jacobson & Pecheone, 1991, pp. 205-206). The project directors believed that "teaching is a complex, multidimensional, contextualized process that cannot be adequately measured either by traditional paper-and-pencil tests or by observational assessments alone" (Jacobson & Pecheone, 1991, p. 206). The program was designed to build a critical link between what teachers knew of the subject matter and their actual teaching practices. The process was to focus on integrating three domains of teaching knowledge: content area knowledge, content/pedagogical knowledge, and knowledge of students. The Connecticut State Department of Education's goal was to integrate the Beginning Educator Support and Training (BEST) Program and the Connecticut Teacher Assessment Center (CONNTAC); thus, making "sound inferences

about a teacher's effectiveness as well as to enrich substantially the quality and usefulness of our feedback to teachers" (Jacobson & Pecheone, 1991, p. 206). In essence, the assessment of teaching performance through combining the two programs was to address competencies that a beginning teacher needed to know and to demonstrate for promoting student learning. Three stages of assessment were developed: essential skills, subject area knowledge, and professional knowledge. The essential skill stage measured whether prospective teacher candidates had reading, writing, and mathematics skills required for teaching. Stage Two, subject area knowledge, ensured that the candidates possessed the content knowledge critical to their subject specialization. The final stage assessed professional knowledge. It incorporated the BEST, which provided support and assessment for beginning teachers, with the Connecticut Competency Instrument (CCI), which measured ten indicators of teacher quality, that were observed in the classroom. Those ten indicators were (Jacobson & Pecheone, 1991, pp. 207-208):

1. The teacher promotes a positive learning environment.
2. The teacher maintains appropriate standards of behavior.
3. The teacher engages the students in meeting the objectives of the lesson.
4. The teacher effectively manages routines and transitions.
5. The teacher creates a structure for learning.
6. The teacher develops the lesson effectively, using appropriate instructional techniques.
7. The teacher presents appropriate content.

8. The teacher uses appropriate questioning techniques.
9. The teacher communicates clearly, using precise language and acceptable oral expressions.
10. The teacher monitors students' understanding of the lesson and adjusts teaching when appropriate.

A pilot study was conducted in November 1989. For the pilot study, assessors were given two days of training to learn about the exercises and the scoring system. Following the training, they scored videotapes of beginning teachers. Candidates demonstrated five tasks on two topics for a total of ten tasks. Inter-rater reliability suggested moderate agreement among raters on the initial independent ratings (overall test reliability was .60); however, when paired ratings that were within one point were included the reliability was approximately .93. "Inter-rater correlations for each rater pair were 1.00, .95, .99, and 1.00, respectively" (Jacobson & Pecheone, 1991, p. 219).

The CONNTAC program was developed to enhance the BEST program by assessing the beginning teacher candidates content/pedagogical knowledge. Two items were learned in the initial stages. One was the important role that subject-area experts had in teacher assessment. The second was the weighty role that scoring had in the process. Through the coordination of the tasks to be measured and the scoring system, an assessment process that was "more focused and aligned with the expected outcomes" (Jacobson & Pecheone, 1991, p. 220) occurred. Tomala and Weinland (1991) reported that "preliminary research suggests that the complexity of the CONNTAC assessment contributes significantly to its

authenticity" (Tomala & Weinland, 1991, p. 202).

Davey (1991) contended that performance assessment should rely on high-inference. "Specifically, it should be capable of evaluating the quality and overall impact of the candidate's performance, rather than simply recording whether behaviors did or did not occur" (Davey, 1991, p. 127). The evaluators needed to be highly trained and carefully guided to achieve reliable results. As a result, the scoring process was based on professional judgment, not mere observational skill. He supported a lower reliability coefficient in exchange for a richer level of behavioral observation. He did, however, contend that acceptable levels of inter-rater reliability could be achieved (i.e., .60 -.80) with the high-inference approach. He viewed the most serious concern regarding performance assessment as the extremely heavy commitment of resources required to develop and implement the program. Davey (1991) believed that the advantages outweighed the disadvantages in that "the performance assessment approach attempts to measure knowledge and abilities *in use* rather than in the abstract" (p. 131).

The evaluation process must provide the district administrators and teaching staff with objective feedback which allows improvement in the district's educational process through the development of its present staff and the hiring improvement of new staff. If the hiring procedures and evaluation procedures do not measure similar characteristics, then one or both instruments must be modified. Hiring quality staff depends upon school personnel being able to assess both candidates and employees using similar, objective criteria.

## Hiring Practices

Currently, school districts have not fully implemented recent research findings and are not using practices that businesses consider routine. Vanderheiden (1982) surveyed personnel administrators to compare recruitment and selection procedures in the United States. He found that schools continued to use similar recruiting and teacher selection processes. All of the surveyed districts used one or more personal interviews, both structured and unstructured. Less than sixteen percent of the schools reported locally-based assessments regarding the validity of teacher selection practices.

The fact is that schools continue to be unique in their selection and evaluation practices. Bolton (1973) identified five points of agreement:

1. The reasons for establishing comprehensive and systematic teacher selection and evaluation procedures are to improve learning conditions for children and to facilitate administrative decisions.
2. Personnel procedures are most effective when they are cooperatively planned by teachers and administrators.
3. All personnel who participate in selection and evaluation should understand their purposes, the nature of the procedures to be followed, and the roles of the other people involved.
4. Training the participants in the specialized task of selection and evaluation increases their effectiveness in performance of those tasks.
5. In order for selection and evaluation procedures to continue to serve their purposes, they must be checked periodically for sources of error.

Prior to the 1970's, school officials discovered difficulties filling many teaching positions. Administrators often found themselves searching for "any"

individual to fill a vacancy. As more teachers became available during the 1970's, administrators were placed in the unfamiliar position of attempting to determine the best candidate. As a result, many administrators struggled with hiring practices. Past hiring practices lacked sophisticated selection criteria offering employers of first-year teachers little more than a 50% chance of success (Diekmann, 1981). Bridges (1986) estimated that 5% of the public classroom teachers were incompetent. According to Fiske (1978), over 50% of the 535 first-year teachers in a school district had failed a mental ability test that was designed to assess high school teachers. Too often the wrong person was hired because of poor selection processes (Heynderickz, 1987; Jensen, 1987).

During the 1970's, school administrators were required to move from a "role of selling the virtues of their school system to a role of trying to sift through the mounds of paper to find a suggestion that a few applicants might be exceedingly better than the rest" (Smith, 1980, p. 312). He cited four factors that negatively affected the ability of administrators to develop quality selection procedures. They were:

1. New, tough affirmative action guidelines.
2. Pressures of coping with rising teacher militancy.
3. Declining enrollment of students in college teacher education programs.
4. Candidates' access to personnel files that allowed them the opportunity to remove or refute documents which were unfavorable.

Essentially, he summed up the problem by stating "the teaching profession has



not established what qualities should be sought or measured in a teacher applicant to distinguish one applicant from another" (Smith, 1980, p. 313).

Vann (1991) identified teacher bargaining contracts and central office priorities (such as, the need for athletic coaches) as major roadblocks to effective hiring decisions. In his district as building principal, he had involvement in the hiring process. He personally screened application letters and resumes; he also required candidates to teach demonstration lessons, submit writing samples, and interview with staff members prior to his ranking and submitting his recommendation to central office administrators. He believed that he had considerable autonomy regarding who was hired, although he did not always get his highest ranked candidate. Whether in a district that promotes site-based management or not, Vann contended that what building principals want was "a chance to influence selection decisions -- and sometimes to point out potential consequences that might have been overlooked -- before hiring decisions are cast in stone" (p. 23).

Ishee (1982) surveyed nineteen suburban school districts in Texas. He found that the larger districts had developed sophisticated selection procedures utilizing a variety of personnel. Candidate assessment was completed using assigned point values to specific criteria. Smaller districts generally did not have a specific evaluation instrument to assess and select candidates.

Brandon (1983) studied the teacher selection process in Tennessee. He mailed questionnaires to superintendents and principals in 145 districts with a return of 109 superintendent and 225 principal responses. Chi square and

one-way ANOVA tests were used to determine statistical significance at the .05 level. His major finding was that superintendents and principals significantly differed in their perceptions concerning teacher selection criteria, regardless of school district size. However, when comparing the superintendents' and principals' perceptions to a national panel of experts for teacher selection, Brandon found a low level of agreement. He concluded that Tennessee administrators needed to better use criteria validated by the national panel.

Jackson (1983) compared opinions regarding the teacher selection criteria among Oklahoma administrators. He received surveys from 501 superintendents, 494 high school principals, 309 middle school principals, and 752 elementary principals. They indicated their perceptions about 58 teacher selection criteria on a five-point Lickert scale. Jackson found significant differences in the perceptions among the superintendents, high school principals, middle school principals, and elementary principals regarding the teacher selection criteria. He found no significant difference among the groups, however, when comparing them according to age, years experience as an administrator, school district classification, and highest degree earned. The criteria considered of little importance among the respondents were race, religious affiliation, number of children, sex, national origin, marital status, socio-economic status, and church participation. Those items ranked highest included dependability, cooperative attitude, enthusiasm, interest in teaching as a career, friendliness, opinions of principals from applicant's previous school, and oral communication skills.

The application letter and/or printed application form were common selection devices used by school administrators prior to the 1980's (Luthy, 1984; Cockrill, 1983; Carrick, 1980). Approximately 91% of the schools in the large metropolitan areas in the United States used application forms in the selection process (Teitelbaum & McLaughlin, 1979). Gerber (1968) studied the effect of a weighted application form on the selection process, finding a correlation of only 0.048 between the weighted application form and success in teaching as measured by principals' ratings.

Carrick (1980) and Leshner and Wade (1972) determined that application letters and forms were the beginning of the hiring process; they allowed district personnel to know who was interested in an opening and let candidates know the hiring steps required by that district. The primary purpose for the application letter was to state an interest in the position, to provide limited information, and to request required application materials. The application letter and forms often eliminated candidates as they failed to complete the required steps, or the forms created negative impressions in the minds of the district personnel who made the hiring decisions.

The resume, as part of the selection process, has been relatively new. Leshner and Wade (1972) found that in most cases the resume was sent unsolicited by the job applicants. Buffie (1980) reported that most districts rarely required an applicant to send a resume. As a result, little effective research regarding its merit in the selection process has occurred. The resume may have provided the district personnel with some additional information; however,

it has been common to find similar information in the application forms and college placement credentials. As a result, the resume duplicated information about the candidate that was generally found in other sources.

College placement offices provide school district personnel with credentials which often include the applicant's transcripts, biographical information, and letters of recommendation. The letters of recommendation can be of two types, confidential and non-confidential. District personnel placed a higher value on those that were confidential; however, all letters of recommendation merely offered "clouded" perceptions. Shelton (1989) found that most people were cautious about putting anything in writing that might later be used against them. He recommended telephone reference checks.

Fuhr (1977) indicated that most student teaching evaluations were vague and highly subjective. The possibility of correlating a grade-point average with actual teaching performance would be difficult at best. Mortaloni (1974) found that the most important letter was from the supervising teacher. Galbo and Diekman (1985) suggested that more weight was given to recommendations from administrators and cooperating teachers than from college supervisors.

Leshner and Wade (1972) sent surveys to 208 district administrators in the Midwest, with 168 surveys returned (81%). They found that credentials were rated extremely high in determining the teacher candidate selected. Credentials were used by 83% of those surveyed. Although district personnel frequently reviewed references when selecting candidates, the references unfortunately were extremely susceptible to bias. Generally, researchers have

found little, if any, correlation existed between recommendations and effective teachers (Hill, 1978; Gerwin, 1974; Arend, 1973; Bolton, 1973; Underwood, 1961). McIntyre (1974) used uniform rating scales in an attempt to improve the predictability of teacher effectiveness. His study also determined that little correlation existed between reference checks and the selection of effective teachers.

Bredeson (1982) researched the effect that the length and favorable tone of the reference letter had on the initial rating and screening of applicants by high school principals. He collected data from 108 randomly selected Midwestern principals. Each principal completed a candidate rating sheet using letters of recommendation and resumes of varying length and favorable tone. A two-way ANOVA was used to analyze the data at the .05 level. He concluded that although there was no statistical significance regarding the length of a letter of recommendation, there was a significant effect on the overall rating of a candidate based on the favorable tone of the reference letter. He also found that age, degree level, and experience factors were not significantly related to a candidate's rating, indicating that items considered to be potential biases in the selection process may not affect the initial candidate rating.

Voss (1983), however, conducted a study to determine if age, teacher position, and the amount of candidate information affected selection decisions. He randomly selected 20 principals from 160 United States public high schools. Each evaluated eight different resumes of a 29-year-old teacher candidate and a 49-year old teacher candidate for the following teaching positions: chemistry,

physical science, health, and physical education. A three-way ANOVA was used to analyze the data. He concluded that the 29-year-old was likely to overall rate above the 49-year-old, especially in the physical education and health areas. He was not able to conclude that providing additional information significantly diminished the effect of age on the candidate rating.

In a related study, Kopetz (1982) researched the effect of non-job related criteria (sex, applied instrument, and the type of institution attended) for first-year teacher applicants seeking employment as instrumental music instructors. He designed "Applications of Employment" that contained the biographical information of four fictitious music education graduates. The "Applications of Employment" and rating forms were sent to 200 randomly selected recruiters. The results of the study indicated a significant difference at the .05 level due to the effects of sex, applied instrument, and the type of institution attended. Recruiters tended to rank male applicants higher, favored trumpet applicants, and ranked graduates from music education institutions higher than those from the performance institutions. Kopetz concluded that recruiters did use non-job related criteria with regard to the employment of instrumental music candidates.

Shanoski and Hranitz (1989) identified grades as the most frequently cited item in the selection of teacher candidates. In a study involving ACT composite scores, Skillett and Tompkins (1984) determined that only in the solid academic areas at the secondary level was there an indication of validity for identifying the appropriate teacher candidate. At the elementary level, the results indicated a random selection of the entire range of abilities rather than

selection of the most academically talented. Young and Elliott (1986) confirmed that the use of grades in the selection process was more important at the secondary level than at the elementary. Despite the use of grades by some in the teacher selection process, Orebaugh (1974) and Thacker (1974) concluded that grades could not predict teacher success.

Place and Kowalski (1993) conducted a study to determine the factors that principals consider to be most important in selecting new staff. They surveyed principals in four Midwest districts. The districts varied from 5,000 to 30,000 in student enrollment. Eighty-one of the eighty-seven principals returned the surveys for a 93 percent return rate. The survey contained forty-six commonly cited characteristics of teachers. Principals were asked to examine simultaneously the importance and assessability for each of the characteristics. Two scales of 1 to 5 were used. In the first scale, 1 represented "not important" and 5 represented "equally extremely important". In the second scale, 1 represented "certain that it can not be assessed" and 5 represented "can be assessed without difficulty". Using ANOVA tests for significance among demographic variables, characteristics that resulted in probability values less than .01 were:

1. Chronological age was found to be significantly less important among elementary principals than among middle school principals.
2. Female principals identified chronological age as significantly less important than did male principals.

3. Principals in one district rated knowledge of multicultural education as less important than did principals in the other three districts; however, it was not the smallest, nor the largest district.
4. The importance assigned to involvement in high school and college activities varied significantly, with elementary principals placing less emphasis on this factor than high school principals.
5. Female principals rated assessing the ability to retain confidentially lower than male principals.

Principals' perceptions showed a high level of agreement about the importance and assessability of the 46 factors. Essentially, principals view the majority of the criteria as important. The most important criteria were also the most difficult to assess. The criteria identified as high in importance but difficult to assess were ability to retain confidentiality, emotional stability, and honesty. Those considered low in importance but easy to assess were age, teaching experience, and involvement in activities. Results from the study indicated a need to "devise assessment procedures that permit administrators to rely on criteria deemed most important" (Place & Kowalski, 1993, p. 300).

Buffie (1980) determined that the interview was the single most important tool in the selection process despite the fact that many (Armstrong, 1988; Cockrill, 1983; Jensen, 1986) questioned its importance. They suggested that the interviewer often asked questions that were too vague and theoretical. The interview was usually more subjective than objective, especially if it was unstructured. They concluded that it was invalid and unreliable. Theodossin (1983) interviewed 409 teacher candidates and found that the correlation between the interview and the success in teaching was +0.12. The correlation



between the interview and successful teaching was practically zero. Luthy (1984) discovered that 95% of the 136 Missouri districts surveyed used interviews. Leshner and Wade (1972) found that district personnel in the Midwest identified the interview as the most important item (88%) in the selection process. Seiferth and Purcell (1979) supported the premise that the interview was an important part of the selection process. Their study discovered that the impression made by the interviewed candidate weighed heavily in the decision made by hiring personnel.

Loehr (1986) determined that the structured interview was the item that principals most often used in determining the best teacher. He was unable to support the principals' decisions based upon statistical data. Ash (1992) encouraged administrators to use a structured ten-minute initial interview for screening a large group of acceptable candidates. Approximately six or seven questions were asked at that time, and candidates who were considerable distances away were interviewed via telephone to save the time and expense of a long trip. The technique could also be used at college job fairs where administrators met many applicants. Ash contended that the quality of teaching must be a high priority and that "if screening gained us a better pool of candidates for our in-depth interviews, we should devote the time necessary to the task" (p. 42). He believed that the ten-minute interview helped the district accomplish its mission of hiring the best applicants available.

Three studies were conducted regarding the use of the Teacher Perceiver Interview (TPI) and its importance within the teacher selection

process. Overman (1982) studied the relationship of the TPI to seven other selection criteria. He obtained results from 271 of 352 administrators surveyed for a 77 percent return. He analyzed the data using the Friedman-Two Way Analysis by Rank and Chi-square statistical procedures. His study determined the TPI as being the most important, followed by the typical interview, references, resume, with the actual application form, transcript, and tests being the least important in the selection process. He found that significantly more users of the TPI (59% versus 41%) validated its ability to identify effective teachers. He concluded that the commitment of time and money for using the TPI in the teacher selection process was supported by the data.

In two other studies, Johannsen (1982) and Cook (1982), the results did not support the use of the TPI as a component of the teacher selection process. Johannsen (1982) studied the data obtained in a research project that involved 40 teachers, 20 hired using the district's traditional selection process and 20 hired using the TPI. The data were analyzed using a t-test. His results found no significant difference between the two teacher selection processes and identified one indicator of effectiveness, i.e., achievement of students in mathematics. Cook (1982) conducted a study to validate the use of the TPI as an instrument predictive of teacher success. Thirty middle school teachers' TPI scores were correlated with the interpersonal relationship behaviors of teachers as measured by FIRO-B. Pearson product moment correlations and t-tests were conducted on the data. No significant results were found, although a positive t-test and correlations among the TPI scores, administrative ratings, and student

ratings did occur.

The Emphasizing More Personalized Attitudes Toward Helping Youth (EMPATHY) Interview was studied by two different researchers. Smedberg (1984) attempted to determine if the EMPATHY Interview was a valid predictor of teacher performance. She studied two groups of teachers, beginning teachers and experienced teachers, hired by a district in 1982-83. A match of demographic variables indicated similarities in seven areas. Both groups were administered the EMPATHY Interview prior to employment and evaluated by principals using the district evaluation instrument for instructional personnel. Also, all participants were observed by trained personnel using the Summative Observation Instrument. Smedberg concluded that the data did not support the EMPATHY Interview as a valid predictor of teacher performance. Correlation coefficients between predictor and assessment instruments ranged between -.10 and +.31.

Smith (1982) conducted a study of the EMPATHY Interview to determine the correlations of scores for a group of teachers. One score obtained during the 1973-74 school year, and the other was a posttest score taken five years later. Student ratings for the teachers were also obtained at both times. He found significant correlations at the .05 level for the EMPATHY pretest, pretest student ratings, and EMPATHY posttests. He also found positive, significant correlations for the pretest student ratings, posttest student ratings, and the EMPATHY posttest scores. Student ratings were found to be the most consistent as measures of perceived teacher competence.

The Omaha Teacher Interview (OTI) was examined in three independent studies to determine whether it was an effective process to identify potentially successful teachers. Solomon (1982) conducted a study of thirty-one elementary and secondary teachers employed in a large suburban school district during the summer of 1979. A correlation was completed using the scores from the OTI and the district's employee evaluation checklist. The checklist divided teacher traits into three categories: 1) Organization and Instruction, 2) Professional Responsibility, and 3) Personal Attributes. An analysis of data found positive correlations, but they were not significant at the .05 level.

Brown (1986) examined the teacher selection process in a school district that used the OTI. A group of ninety-six teachers were divided into two groups, those hired using the OTI as a part of selection process and those hired without using the OTI. The individual scores on the district's teacher evaluation instrument were compared using t-tests. No statistically significant difference between the two groups was found at the .01 level ( $p < .021$ ).

McGarity (1987) studied a sample of eighty-one teachers hired during the 1982-83 school year in a school district. Using the results from the district's evaluation instrument, he conducted ANOVA tests to determine whether significant differences existed between the groups by teaching level. He found no significant differences among the participants, concluding that the OTI did not adequately distinguish and identify top-caliber teachers. None of the studies were able to determine a significant difference in identifying teacher

candidates who would be outstanding teachers by using the OTI as a component of the selection process.

Despite the heavy use of interviews in the selection process, its validity and reliability is questionable. Cherrington (1987) stated:

Unfortunately, research evidence does not support the value of employment interviews. Interviews are conducted by individuals who have different orientations, different levels of competence, and different biases. The interview process is generally not consistent--each interview is conducted differently depending on what is said--and the evaluations of the interviewer are essentially random observations. Although interviewers may believe their conclusions are the most important step in the decision process, research shows that interview data are generally neither reliable nor valid (p. 78).

Al-Rubaiy (1993) advocated a five-step approach to hiring. Those steps included recruiting, initial screening, application submission and screening, team interviewing, and decision-making. The district sought applicants through newspaper advertisements and college/university placement offices. Upon receiving that list of applicants, an initial screening was completed to identify those individuals who met the district's needs. The selected candidates completed an application and provided "a short, handwritten writing sample, transcripts, and a copy of a current teaching certificate" (Al-Rubaiy, 1993, p. 21). Candidates selected for interviews met with teams consisting of central office administrators, principals, teachers, and parents. Following each interview, team members completed an assessment of the applicants, identifying whether the individual was highly qualified, qualified with reservations, or unacceptable. The final decision was made only after a team debriefing, at which time team

members listed those candidates they believed to be highly qualified and the reasons for their choices. Following reference checking, a decision was made concerning to offer the contract.

The emergence of work samples in the selection process may prove promising. Businesses have been using work samples to improve the hiring process (Jensen, 1986) for some time. Support for the use of work samples stems from logic. The selection process should be based upon the district's expectations of a teacher after being hired. Pounder (1989) suggested that assessing teacher candidates on skills that mirror those required on the job may increase the predictive validity of selection decisions, thus, "making teacher selection an effective, viable, and legally defensible means to increase the quality of teachers in schools and in the profession" (p. 149). Title (1995) cited three reasons to expand the hiring process to include authentic assessment :

1. The belief that the traditional interview method of screening candidates could be improved.
2. Teachers are hired to teach, not interview.
3. The political realities require the need to have clear, agreed-upon criteria.

The lack of a strong research base, supporting the interview as appropriate in identifying the best applicant, led Title (1995) to seek alternatives to the traditional hiring process. In essence, the assessment of candidates should address the major functions of the position. "When the selection process involves a portfolio assessment, evaluation of student work, planning lessons, and actual teaching with real children, then it is much more difficult for people to

claim successfully that the final selection was made on the basis of politics and not merit" (Title, 1995, p. 6).

Only one study was found that examined the selection and evaluation policies and instruments used in school districts. Ackerson (1977) studied thirty-five school districts in Illinois to determine if there were similarities between the criteria used in teacher selection and teacher evaluation. The schools chosen for the study were those districts with a student population of 5,000 or more, excluding the Chicago Public Schools. Of the thirty-six districts identified, all but one participated. (It was unable to participate due to its district policies.) Ackerson (1977) identified only fourteen of the thirty-five districts who actually used selection and evaluation policies and instruments. He concluded that the criteria used to screen applicants was not necessarily used to evaluate staff later. For example, elements such as respect for students, providing for individual differences, or maintaining a congenial learning atmosphere were not identified before the teachers were hired, yet he contended these criteria were deemed important in the teacher evaluation process, hence, they could have been and should have been investigated with a thorough reference check of former employers, professors or student teacher supervisors.

Criteria used must be evaluated to see whether it is producing the individual best suited to the system. If this criteria is similar to that used in the appraisal of teachers on the job, the district should also be able to assess its evaluation process... The inclusion of the same basic criteria in both selection process and in the evaluation process should help all individuals in the system know what is expected from them on the job, and provide personnel with opportunities to grow in directions that will satisfy individual needs and system expectations (Ackerson, 1977, p. 80).

Cockrill (1983) discovered that Texas personnel specialists reported six items as being essential to the teacher selection process:

1. Interest in children
2. Emotional stability
3. Ability to communicate with pupils
4. Ability to get along with pupils
5. Ability to maintain classroom discipline
6. Enthusiasm for teaching

He, however, determined that the personnel specialists relied more heavily on subjective criteria and procedures rather than seeking objective, factual data. The greatest value was placed on the oral interview.

Wise, Darling-Hammond, and Berry (1987) identified the observation of teaching performance as "probably the most reliable and valid assessment available to school district administrators" (p. 67). School district personnel could, with minimal effort and costs, incorporate the TDL into the hiring process, directly linking it with the district's expectations of teachers already employed. "The financial costs of this organizational commitment to collect the most valid information are modest for a school district... At most, the district must hire substitutes to cover the classes of those teachers involved in the selection process..." (Wise, Darling-Hammond, & Berry, 1987, p. 65).

Including an observation of the candidate's teaching ability in the teacher selection process was appropriate for identifying the individuals who would be able to perform in the classroom according to the district's standards (Fitzpatrick



& Morrison, 1971). Bolton (1971) identified that the primary advantage for using simulation was doing it in a controlled environment which could increase the measurement reliability. Reilly and Chao (1982) found in their study that the validity of work samples ranged from a correlation of 0.21 to 0.44. Two traditional components in the selection process, interviews and reference checks, had only correlations of 0.19 and 0.18, respectively. The need to view a candidate teaching a micro-lesson was suggested by some authors (Nicholson & McInerney, 1988; Bredeson, 1985); however, empirical data was lacking. Seventy-six percent of the secondary administrators "indicated they would be interested in seeing a video-taped lesson" (Braun, Willems, Brown, & Green, 1987, p. 48).

Nalley (1971) surveyed twenty-one large metropolitan school districts and found that 70% rarely, if ever, observed candidates. Luthy (1984), however, found that one-third of the 136 Missouri districts he surveyed used observations as a part of the selection process. Caliendo (1986) suggested that applicants should submit a video recording of a teaching segment which would provide firsthand insight into one's teaching abilities. He hired fifteen teachers using observation as a part of the selection process. At the end of the first year of teaching, all had received exceptionally high performance ratings; however, no statistical tests were conducted to determine if the evaluations were significantly different from staff previously hired without using a TDL.

Callaghan (1982) studied a teacher selection procedure that used a wide variety of screening techniques, on-the-job observations, and rankings of

teacher classroom performance. Eighty-one teachers hired in a Chicago magnet school during the 1971-1974 school years participated in the study. The selection process used a battery of personality and attitudinal inventories (Minnesota Teacher Attitude Inventory, Rokeach Dogmatism Scale, Young Change Inventory, Teacher Practices Questionnaire, and Verbal Facility Test). In addition, candidates were required to complete an essay about their motivation to teach. Administrators conducted pre-selection classroom observations. Interview teams included parents and community members. The pre-selection components that showed significant relationships at the .05 level included the personality and attitudinal inventories and the ratings by the parents and community members.

LeTendre (1989) surveyed 106 California school districts with 90 responding (an 85% return). She identified five different types of classroom observations:

1. A mock lesson taught to a panel of adults
2. A demonstration lesson taught to students
3. An observation of teachers in their own classrooms
4. Substituting
5. Student teaching

Of those responding, 24 districts (or 27%) indicated that they used at least one of the classroom observations listed above. Fifteen school districts were selected for closer study. LeTendre conducted interviews with the district personnel to determine the steps involved in the selection process and what

importance district personnel gave to the work sample of candidates within the selection process. District personnel indicated that they believed they were able to make more accurate selection decisions and that better quality candidates were hired as a result of observing the individuals teach. LeTendre recommended that a formal job analysis be completed prior to designing any observation exercise and suggested that the observation process be standardized as much as possible. Improvement of the predictive validity and fairness of the TDL observation requires formal training for the evaluators, using more than one evaluator, and allowing candidates to display the skills, knowledge, and values deemed important for success in the position.

Browne's (1972) study involved a comparison of the traditional teacher selection process (review of credentials and personal interview) and a classroom-observation teacher selection process. He videotaped five first-grade teachers from the same school district during a structured interview and made a thirty-minute videotape of each teacher presenting a reading lesson, a math lesson, and show and tell. Samples of three activities were videotaped randomly without the teacher or students necessarily knowing that the camera was operating. The final thirty-minute tape contained approximately the same time allotments of the three activities for each teacher. Twelve practicing principals were randomly selected as the judges and divided into two groups of six each. One group ranked the candidates after reviewing the credentials and again after viewing the interview videotape. The second group ranked the candidates after reviewing the credentials and again after viewing the teaching

video. Two weeks later, Browne had each group rank the candidates after viewing the videotape they had not seen earlier. Browne did not find a statistically significant difference regarding the viewing of the videotapes as a part of the selection process. In fact, he was not able to obtain significant agreement among the judges for either of the selection processes. He recommended that the use of videotaped lessons become a part of the selection process to possibly improve the validity of the practice. Browne suggested that if the practice of using videotapes was to become more prevalent, administrators would become better at using videotapes in the selection process. As a result, the concept would become a more appropriate tool for teacher selection.

Stone (1972) studied the effect of live teaching observations and videotaped teaching observations on teacher selection decisions made by district administrators. Twenty of the twenty-six judges were doctoral students in educational administration. The other six were practicing administrators or former administrators who were faculty members in the College of Education at the University of Oklahoma. The judges were divided into three groups, each ranking six teacher candidates in order of their suitability for selection. The control group used traditional selection methods while one of the two experimental groups observed a live teaching demonstration in addition to the traditional selection methods. The other experimental group watched a videotape of the teaching demonstration and used the traditional methods. Using Kendall's coefficient of concordance and Spearman's rank correlation

coefficient, Stone found the live observation of the teaching demonstration to be significant at the .05 level. The viewing of the videotape was not significant in determining the suitability of teacher candidates. Stone noted that with more use the method might prove to be effective in time.

### Summary

Despite the importance of the teacher selection process in providing quality student learning opportunities, only a minimal amount of research regarding the use of a TDL has emerged involving the identification of more appropriate selection procedures than those presently in use (Browne, 1972; Stone, 1972; LeTendre, 1989). Administrators generally used limited information, much of which was obtained from the perceptions and opinions of others found in application materials, personal interviews, and telephone reference checks (Moore, 1987; Kirkland, 1979; Ackerson, 1977; Wasicsko, 1977; Merritt, 1971).

If district personnel are to improve upon the selection of quality teachers, then it is imperative that they use the knowledge available, changing to a process based upon objectivity and eliminating as many biases as possible. When districts set goals to hire quality instructional staffs, it is critical that they take the necessary steps to insure the best candidates are hired. This can occur only when district personnel complete all of the available criteria within the process as determined by that district. This change in selecting staff requires that school administrators begin to use the new data suggesting

modifications in business-as-usual, which were recommended by LeTendre (1989), Browne (1972), Stone (1972), and others. It requires that candidate screening be based upon the objective teaching criteria used within the district's evaluation system which directly affects student performance. It also requires that administrators seek out information from a variety of sources, one of which may be the actual observance of a candidate demonstrating teaching skills.

## Chapter 3

### DESIGN OF THE STUDY

This study was designed to determine whether the inclusion of a teacher demonstration lesson as part of the teacher selection process helped identify teachers who would be more likely to achieve higher performance ratings after they were hired. The research question that guided this study was:

Does the distribution of ratings on the SER for teachers employed using the TDL differ from those employed without using the TDL?

The following details the independent variables, dependent variable, population, rights of human subjects, instrumentation, data collection, treatment of the data, null hypothesis, and summary.

This was a study of teachers employed in a Midwestern school district between 1980 and 1988 who remained employed in that district during the 1990-91, 1991-92, or 1992-93 school year as classroom teachers. Those employed between 1980 and 1983 were hired using traditional selection procedures. The second group of teachers was hired between 1985 and 1988 using the TDL as a component of the selection process. A comparison of the two groups' individual teacher's most recent evaluation ratings was then used to determine if differences existed between the evaluations for the two groups. Between 1983 and 1985, administrators modified both the selection process and the SER instrument. The selection process was changed to include the TDL as a part of the interview, and the SER instrument was revised as a part of the School Improvement Model (SIM) Project. It was during this two-year time

period that district administrators were able to practice and develop their observation and evaluation skills. Beginning with the 1985 school year, all teacher applicants interviewed by the district administrators were required to present a TDL as a component of the selection process.

<u>COMPARISON GROUPS</u>		<u>EVALUATION COMPLETED</u>		
		1990-91	1991-92	1992-93
District with TDL as a part of selection process (N=101)				
A-1	Teachers hired between 1980-83 without TDL (N=53)	X	X	X
A-2	Teachers hired between 1985-88 with TDL (N=48)	X	X	X

---

Figure 1. Counts for Non-TDL and TDL Teacher Groups

---

The district had been involved in using the TDL for the past seven years and the data were taken from evaluations that had been completed before the study began; therefore, the study used an ex-post facto research design.

#### Independent Variables

The independent variables used in this study were the selection procedures prior to 1983 and those used after 1985. Detailed below are the procedures used by the district as outlined by the district's personnel director in interviews with the researcher and from the district's procedural manual. Prior to 1983, candidates submitted written application materials consisting of an application letter, application form, resume, transcripts, and credentials to the district personnel director. The building principal submitted a written summary of required skills and desired characteristics as selection guidelines, which was developed through input from building staff and, in some cases, parents.



District personnel reviewed the application materials based upon those guidelines to select candidates to be interviewed. The building principal then conducted a traditional interview with these candidates. This interview provided information specific to the assignment, candidate, and the school.

The district also used a structured interview from Ventures in Excellence, a private educational research company located in Lincoln, Nebraska. This company developed a research-based interview process that focused on the attributes of successful teachers. The district personnel who conducted the structured interview had successfully completed the comprehensive training provided by the company. In most interviews the building staff and parents were involved in the interview process. During this time, the building principal observed how each candidate developed group rapport. Staff and parents provided input as to their perceptions of how the candidate might relate to colleagues, parents, and students and the type of role model that the person might be for students. Using this process, the principal was able to validate conclusions developed from the structured interview. Reference checking verified the information supplied by the candidate and obtained during the interview process. The principal then submitted to the personnel director a written recommendation identifying the candidate to be offered the position, based upon the specific skills and characteristics of that individual in relationship to the written list of selection guidelines.

In 1983, district personnel began, on a pilot basis, to include the TDL as a component of the interview process on a pilot basis. Some staff were then

hired based on this additional information. During 1983-85, building-level administrators practiced the evaluation skills prior to being required to use them on a permanent basis for the evaluation of staff and candidates. Beginning with the 1985-86 school year, all candidates hired were required to perform a TDL as a part of the selection process.

The school district's rationale for the TDL was that interviews should not be limited to assessing an individual's skills only by asking others to describe how the person teaches or by asking the candidate to describe how he/she teaches (through direct or situational questions). Instead, they thought that district administrators should make efforts to gain first-hand information by having candidates demonstrate, rather than merely describe their teaching skills where candidates may portray themselves as more successful than was actually the case.

All candidates invited to interview for a teaching position in the school district were sent an explanation of the TDL prior to the interview. Candidates were instructed to present a 15-20 minute teaching lesson that covered any piece of content or skill appropriate for the designated grade level(s) in a particular curriculum area determined by the district. The individual taught the lesson to a group of five to six volunteer students who, when possible, were at the same grade level for which the candidate was interviewing. These volunteers were treated as if they were students in a regular, on-going class. The building principal, appropriate central office consultants, and the personnel director observed the teaching vignette. Each candidate provided a written

lesson plan to the observers prior to the presentation. The observers took an anecdotal script and analyzed the lesson using specific and general criteria.

The specific criteria included:

1. Was there an identifiable objective and was it being taught?
2. Was there an appropriate content decision and did it fit the time allotment?
3. Was evidence sought from students concerning to what degree the objective was achieved?

The general criteria included:

1. Were the basic principles of motivation and reinforcement used?
2. Were appropriate examples/models used?
3. Were other appropriate principles of learning used?

#### Dependent Variable

The dependent variable in this study was the summative evaluation instrument used by the school district for each teacher in the study. The SER used by the district administrators was modified during the 1983-84 and 1984-85 school years with the new instruments fully implemented for the first time during the 1985-86 school year.

All of the evaluation results used in this study were of tenured teachers. The rationale for using only tenured teachers was due to the modification of the evaluation instruments. Since the evaluation instrument was changed after the evaluations were completed for the first group during their probationary period, it was not possible to use the evaluations completed for each group during their probationary period. Building-level administrators completed an in-depth

evaluation for each tenured teacher on a three-year rotating cycle. The formal evaluation process included four observations. Each observation was followed by a post-observation conference with a minimum of two of the observations being preceded by a pre-observation conference.

The district continued to collect data concerning the administrative ratings to determine whether the instrument was reliable and valid. The personnel director shared the results with the teacher education association. Not a single grievance was filed against the district questioning the validity or inter-rater reliability of the instrument or the evaluators. A primary outcome for the district was the development of a research-based, valid instrument for the evaluation of teaching staff. The criteria listed in the performance areas of the SER were determined while the district was a participant in the Iowa State University SIM research project. Administrators involved with the SIM project were trained to become more effective evaluators. They participated in teacher evaluation sessions, during which time they were provided machine-scored, computer-based analysis indicating how each person evaluated teachers and the inter-rater reliability of those evaluations (Manatt, 1982).

The district's instrument (Appendix A) contained the ratings of "performance exceeds district standards", "performance meets district standards", and "performance does not meet district standards." The procedures required that an evaluator submit evidence to document the marking of "performance exceeds district standards" or "performance does not meet district standards." Evaluators were encouraged to add comments on the

instrument for "performance meet district standards".

### Population

The teacher evaluations were of classroom teachers employed in a K-12 public school district located in the Midwest between 1980-83 and 1985-88 and who received a summative evaluation during the school years of 1990-93. A total of 101 individuals hired during that time period remained employed as classroom teachers. Fifty-three individuals were hired between 1980 and 1983 when the TDL was not part of the selection process. Between 1985 and 1988, the district hired forty-eight individuals using the TDL as a part of the interview process. The district's seniority lists from 1990 to 1993 identified the individuals, the date of selection, and current assignment. Since individuals in the study were those who remained as classroom teachers between 1990 and 1993 and each participant's latest evaluation used, there was no loss of subjects. Therefore, mortality should not be considered a threat to the study.

The district was located in a city of approximately 60,000 people with an estimated one-half million people living in an adjacent metropolitan area. The area was representative of the midwestern region, containing a mixture of blue collar and white collar workers. The locale had a two-year community college and three four-year colleges or universities (one public and two private) as well as several vocational and technical schools. The K-12 student population was about 10,150 with approximately 690 certified staff members. The administrators who evaluated teachers possessed the appropriate licenses and endorsements.

### Rights of Human Subjects

The school district's superintendent and district personnel director granted permission to conduct the study and provided the information regarding individual teacher's evaluation data from the district's database. This information was provided in a format that assured confidentiality and made it impossible to attach names to the data. Permission was requested and granted from the Drake University Human Subjects Committee prior to conducting the study. All correspondence between the school district's personnel and the researcher stressed that confidentiality of the data would be maintained and that the district would not be identified by name.

### Instrumentation

Data were collected from two sources. The first was the district's seniority lists for the 1990-91, 1991-92, and 1992-93 school years. The second was the district's summative evaluations obtained between 1990 and 1993. The school district was a participant in the School Improvement Model (SIM) under the direction of Dr. Richard Manatt, Iowa State University. The SIM project provided the participants with appropriate training and practice concerning the teacher performance areas, criteria response modes, and standards (Manatt, 1982). The primary purpose of the SIM project was to improve teacher performance and administrative evaluation skills with the ultimate goal of improved student learning. The district's steering committee developed the instrument based upon the research on effective teaching behavior. Since the district did not modify its summative evaluation instrument during the time of the study, a threat

to the instrumentation used in the study does not pose a problem. With the instruments developed under the auspices of Dr. Manatt, the evaluation instruments should be considered to have face validity.

Kelly (1989) identified four behavioral variables through a review of the literature that influenced and promoted effective teaching. These variables were a strong sense of the educational and instructional goals, student centeredness, use of a variety of teaching methodologies, and appropriate and evaluative feedback. In comparing 74% of the SERs used in Iowa schools during the 1987-88 school year, Kelly determined that items addressing the critical issues represented 56% of the total number of items. This was significantly above the 50% level; however, it was still considered relatively weak. Kelly concluded that SERs could be used more effectively by designing them with four variables of effective teaching in mind. The SER for this study's district contained sixteen of the eighteen items or 89% relating to the four critical issues. This would indicate that the instrument was a valid measurement of effective teaching behavior.

The threats of testing and statistical regression required that a pretest be taken that would influence scores on a later test. Since the individuals had been evaluated using the same instrument in the past, some may argue that it was a form of pretesting. The SER, however, was considered the documentation of the teacher's performance during that evaluation cycle. Although teachers and administrators were knowledgeable of the expectations of the district and the standards to meet those expectations, the previous

evaluations were not considered pretesting. Therefore, testing and statistical regressions were not considered threats to the study.

#### Data Collection

The researcher corresponded with the district's personnel director regarding the possibility of studying the district's teacher selection process and evaluation data to determine if the inclusion of a TDL as part of the process helped identify teachers who would be more likely to achieve higher performance ratings on the SER. Following discussion of the study's intent, the district superintendent and personnel director granted permission to conduct the study and gave assurances that they would provide the needed data in a manner that maintained confidentiality for the staff. The proposal was then submitted to the Drake University's Human Subjects Committee who granted permission to conduct the study.

The personnel director provided this researcher the district's seniority lists for the 1990-91, 1991-92, and 1992-93 school years. These lists contained the names of each teacher employed for that year, the subject/grade level taught, and the year hired. The researcher identified from the seniority lists those individuals presently teaching who were hired during the two time periods being studied. The names of the individuals in the two groups were submitted to the personnel director. To ensure confidentiality, district personnel provided the evaluation data for each subject within the two groups in a coded format.

#### Treatment of the Data

For this study, a Chi square test with nonparametric statistics was used to



determine if there were significant differences between the variables at the .05 level. The use of a Chi square test made the assumption that the scores (the number of times an administrator rated a teacher in that category) were independent, and that there was no relationship between the frequencies or the method of obtaining those frequencies. It also permitted the measure of the magnitude between the differences of observed frequencies (Comparison Group A-2) and expected frequencies (Comparison Group A-1), regardless of the type of data obtained (Borg & Gall, 1989; and Ferguson & Takane, 1989).

Within this study, there was no reason to believe that the scores were dependent of each other or in the method of obtaining them. Statistical regression would have occurred whenever pretesting was used. Since there was no pretest involved, statistical regression was not a threat in the study.

#### Null Hypothesis

The following hypothesis was stated in the null form and evaluated at the .05 level of significance.

Null Hypothesis: There were no differences between teacher evaluation ratings as measured by the Summative Evaluation Report (SER) for teachers selected using a Teacher Demonstration Lesson as part of the selection process and those who did not have TDL as a part of their selection process.

#### Summary

The purpose of this chapter was to discuss the research methods and procedures used in the study. The discussion reviewed the research design, independent variables, dependent variables, population, rights of human

subjects, instrumentation, data collection procedures, treatment of the data, null hypothesis, and summary. The following chapter will provide a presentation and analysis of the data collected in the study.

## Chapter 4

### ANALYSIS OF THE DATA

While many believe that the hiring of highly qualified teachers is one of the most important responsibilities of school district administrators (Vann, 1989; Wise, Darling-Hammond, & Berry, 1987; Herman & Stephens, 1987; Moore, 1987; Caliendo, 1986; Cardozier, 1985; and Smith, 1980), little research has been conducted to determine if traditional paper screening and interview practices are more able to identify effective teachers than requiring candidates to demonstrate their teaching skills (LeTendre, 1989; Ackerson, 1977; Browne, 1972; and Stone, 1972). There are research projects that have studied alternative approaches beyond the traditional practices, however, most have shown inconclusive results (McGarity, 1987; Brown, 1986; Smedberg, 1984; Callaghan, 1982; Cook, 1982; and Ishee, 1982).

The purpose of this study was to determine if the inclusion of a teacher demonstration lesson as part of the teacher selection process helped identify teachers who would be more likely to achieve higher performance ratings after they were hired. The research question that guided this study was:

Does the distribution of ratings on the SER for teachers employed using the TDL differ from those employed without using the TDL?

The study reviewed the evaluation data of 101 teachers employed in a Midwest school district. Forty-eight of those teachers were hired between 1980 and 1983 without the TDL as a part of the selection procedures. Between 1983

and 1985, the district modified its selection procedures to include the TDL as a part of the interview. During that time frame, district administrators were provided training concerning evaluation procedures and were able to practice observing teacher candidates without being required to use the TDL as a part of the selection process. In addition, the district participated in the Iowa State University School Improvement Project under the direction of Dr. Richard Manatt (Manatt, 1982). The district personnel selected fifty-three teachers between 1985 and 1988 using the TDL as a part of the selection procedure.

A Chi square test was performed on the data from each category to determine if a relationship existed between the two groups (Non-TDL and TDL). The Chi square test is a nonparametric statistical test that permits the examination of discrete data (Borg & Gall, 1989; and Ferguson & Takane, 1989) which is how the data for this study were reported. The test is the most frequently used procedure for this type of data, because it permits the comparison of the frequency counts actually obtained with a set of theoretical frequencies (Borg & Gall, 1989; and Ferguson & Takane, 1989). The use of the Chi square made the assumption that the scores (the number of times an administrator rated a teacher in that category) were independent, and that there was no relationship between the frequencies or the method of obtaining those frequencies. Within this study, there was no reason to believe that the scores were dependent of each other or in the method of obtaining them.

### Item Analysis and Results

Table 1 shows the calculation of Chi square in comparing the observed and expected evaluation frequency counts of the ratings on the SER for the Non-TDL and TDL teacher groups. An analysis of the data for each of the four performance areas was completed as well as for the composite ratings. None of the areas demonstrated a significant difference between the observed and expected frequency counts for the administrator ratings.

In one area, Classroom Management, a Chi square value of zero was obtained, indicating no difference between the expected and observed frequency count. In a second area, Professional Responsibilities, the Chi square value obtained was minimal. A third area, Effective Interpersonal Relations, had a Chi square value of only 1.9. The Chi square results for these areas would indicate little, if any, difference existed between the TDL and Non-TDL groups.

A minimal difference, although not significant, occurred between the two groups in the area of Effective Interpersonal Relations. The area of Productive Teaching Techniques and the Composite Score produced the largest variance with Chi squares scores of 4.2 and 3.6 respectively. A Chi square value of 5.99 would have yielded a significant value at the 0.05 level and a value of 9.21 at the 0.01 level. Since none of the areas obtained a significant value at either level, the null hypothesis was accepted.

---

Table 1  
Chi-Square Test of Observed and Expected Evaluation Frequency Counts for  
Non-TDL and TDL Teacher Groups

Group	Observed		Expected		*
Performance Area I: Productive Teaching Techniques					
	Non-TDL	TDL	Non-TDL	TDL	
Does Not Meet	3	0	1.6	1.4	
Meets	314	275	309.1	279.4	
Exceeds	54	61	60.3	54.7	4.2
Performance Area II: Classroom Management					
	Non-TDL	TDL	Non-TDL	TDL	
Does Not Meet	0	0	0	0	
Meets	183	166	183.1	165.9	
Exceeds	29	26	28.9	26.1	0
Performance Area III: Effective Interpersonal Relations					
	Non-TDL	TDL	Non-TDL	TDL	
Does Not Meet	1	0	0.5	0.5	
Meets	231	216	234.6	212.4	
Exceeds	33	24	29.9	27.1	1.9
Performance Area IV: Professional Responsibilities					
	Non-TDL	TDL	Non-TDL	TDL	
Does Not Meet	0	0	0	0	
Meets	86	81	87.6	79.4	
Exceeds	20	15	18.4	16.6	0.3
Composite Score					
	Non-TDL	TDL	Non-TDL	TDL	
Does Not Meet	4	0	2.1	1.9	
Meets	814	738	814.4	737.6	
Exceeds	136	126	137.5	124.5	3.6

\* Indicates results significant beyond the .05 level

## Summary

This chapter reviewed the analysis of the data. The number of participants involved in the study and an examination of the hypothesis was presented. None of the Chi square values for the data on the four components of the evaluation instrument and the composite score yielded a significant difference; therefore, the null hypothesis was accepted. The final chapter provides a discussion of the research project, conclusions reached, implications derived from the study, and recommendations for further study.

## Chapter 5

### SUMMARY, FINDINGS, IMPLICATIONS, and RECOMMENDATIONS

#### Summary

The literature review conducted for this study indicated teachers have a significant affect on student achievement (Title, 1995; Ross, 1991; Cox, 1981; Carrick, 1980), and that school administrators can greatly impact on instructional programs by selecting excellent teachers (Herman & Stephens, 1987; Boyles & Engel, 1986; Bredeson, 1985; Cardozier, 1985). A limited amount of research concerning the use of teacher demonstration lessons (TDL) within the selection process indicated a need to further study teacher selection procedures. The purpose of this study was to examine the potential use of a TDL in one school district's teacher selection process by determining if there were differences in the evaluations of teachers hired using the TDL and those hired without using the TDL.

The study consisted of 101 teachers employed in a school district, 53 employed between 1980-83 without using the TDL and 48 employed between 1985-88 using the TDL. Since the district had modified its SER between 1983 and 1985, it was not feasible to use the evaluation ratings for the teachers during their probationary period. Therefore, the evaluation ratings used were from 1990-93. A Chi square test was used to determine the significant probability of a null hypothesis. The data were analyzed according to four performance areas and an overall evaluation score on the district's SER. Each of these areas failed to determine significance probability at the .05 level; thus,



the data failed to reject the null hypothesis.

### Findings

The results of the study did not generate evidence that the use of the TDL within the selection process provided the district with better assessment of the candidates. Thereby, making the findings inconclusive as to whether better qualified teachers were hired by the district as a result of including the TDL within the selection process. The results of the study also indicated that the evaluation instrument lacked the required discrimination factor, having only three standards by which to measure teacher performance. The two extremes required administrators to document evidence for the assessment. This requirement may have influenced their decision to mark marginal performances in the "meets standards" category rather than appropriately marking and documenting the deficiencies. Similarly, administrators may have limited their ratings of exemplary performance for similar reasons.

### Implications

The eighties and nineties have been filled with new technology in the classrooms. Students are "surfing the Net" to access data not previously available to students and communicating with students in other worlds and world leaders via e-mail. Most researches continue to determine that the greatest impact on student learning remains the classroom teacher. Current evaluation procedures often lack the discrimination required to identify superior teaching performance. Recent work in this area has shown some promise. Davey (1991) supported the use of high-inference assessment. He contended

that evaluations should assess the quality and overall impact of the person's performance, rather than merely record what did or did not occur. Quality assessment of teaching performance will require considerable time on the part of administrators for staff development and in collecting and analyzing the data. Districts must be willing to provide the required training for administrators to learn and practice authentic assessment procedures. Teacher evaluation could be improved through the use of more appropriate, authentic assessment procedures that provide the teacher with objective, diagnostic feedback which encourages the individual to set appropriate, professional growth plans. Efforts must be made to accurately assess teacher performance in the classroom. The evaluations must be based upon appropriate teacher behaviors that the district has identified as critical to promote student learning. Administrators must be provided appropriate training to learn how to assess and document teacher performance within the classroom setting. They must also be allowed the time required to complete the evaluation process and demonstrate a commitment to allocate their efforts in this process. Evaluations should be developed that require written documentation for all assessments. If scales must be used, they should be lengthened to increase the ability to discriminate and accurately assess a teacher's ability rather than grouping everyone as adequate or "meets standards." Student learning can not improve unless continual improvement is sought by the classroom teacher. With the wide range of abilities and needs of today's students, classroom teachers must use the available knowledge base and learn skills to provide an array of learning activities required in today's

classrooms.

Past selection procedures lacked sophisticated selection criteria and, at best, gave employers of first-year teachers little more than a fifty percent chance of success (Diekmann, 1981). The traditional interview may indicate that some have the knowledge and experience; however, additional objective data can and should be collected. Although the interview has been the most used process for selecting teachers, its effectiveness has not been substantiated. The need to revise current practices and implement a reliable, valid, and practical method of assessing and predicting teaching success is vital, if selection procedures are to improve. All too often, administrators fail to use important selection information because it is not readily available nor is it considered essential (Wise, Darling-Hammond, & Berry, 1987; Ackerson, 1977; and Brooks, 1967). Effective administrators need to learn to utilize all available components of the teacher selection process. Teacher candidates should be assessed on skills that mirror those required on the job. The development of a written selection process that includes the TDL may increase the predictive validity of selection decisions. At a time when school finances are limited and closely scrutinized, it is imperative that school administrators wisely allocate the district funds. Administrators need to be aware of the fact that reliable, cost-effective tools are available for their use. Wise, Darling-Hammond, and Berry (1987) stated that the observation of a teaching performance was a valid, reliable, and cost effective assessment method. The cost factor need not be a deterrent in implementing this procedure.

Classroom observation indicates the extent to which a candidate reveals appropriate and adequate knowledge, interactive skills, and teaching strategies. In effect, professional references can provide appraisals of past performance and classroom observation appraisals of current performance. Because past and current performance are the best predictors of future performance, these mechanisms may provide the most reliable and valid assessment of how effectively candidates will teach. (Wise, Darling-Hammond, & Berry, 1987, p. 64)

A review of the literature indicated that valid, objective data is available to administrators regarding potential teachers; however, many administrators fail to access that information. Instead, they tend to rely on subjective criteria. Administrators need to make a commitment to direct more of their time to teacher selection if it is to be improved. It requires that administrators take time to develop a written teacher selection process that emphasizes the use of objective selection criteria and then take the time to use that selection process. It would appear that teacher selection could be improved if administrators would take time to collect the available objective data and then use that information in making sound hiring decisions. Teacher selection remains a primary function of school districts. Administrators will have the greatest impact on student learning through the improvement of their selection process, which must include the objective assessment of criteria based upon the district's evaluation standards.

Finally, administrators need to develop and follow specific steps within the selection process, regardless of the procedures used. Some steps are critical if appropriate selection decisions are to be made. A critical step that often appears to be ignored in many districts is the development of a written list of the essential and desired qualifications required of applicants prior to

advertising the vacancy. Without it, administrators easily rely on subjective data and the "good old boy or girl network", two aspects that have no place in making effective, appropriate hiring decisions.

### Recommendations

Future studies regarding teacher selection procedures involving the demonstration lesson should look at whether better candidates are selected through the study of the evaluations for probationary teachers. The evaluations may provide data that permits the opportunity to identify the individual strengths and weaknesses of staff as administrators closely scrutinize the teaching behaviors during that probationary period. The evaluation ratings may provide administrators with objective data as to whether better applicants were hired as a result of the selection process.

Additionally, a study regarding the number of probationary teachers who continue to be employed as tenured staff should be considered. The data from such a study could be valuable in determining whether quality teachers were hired through the selection process. If the number of probationary teachers granted tenure is different for those hired using the TDL as compared to those employed without its use, then the TDL may be appropriate as a component of the selection process. As districts strive to improve the quality of student learning opportunities through the employment of the best teachers, the use of TDL as a part of the selection process may increase. If district personnel become more accustomed to the use of the TDL, it could improve the validity of the practice. As a result, the concept would become a more appropriate tool within

the teacher selection process.

Finally, a qualitative study regarding the perceptions of those involved in the selection process should be completed. The study could identify how administrators, staff, parents, students, and applicants view the process. The quality of the selection process may be improved through such a study. If each group perceives the TDL as an important component of the selection process and better decisions are believed to be made as a result of including it, then the quality of teachers and presumably student learning may be enhanced.

## REFERENCES

- Ackerson, G. L. (1977). A comparison of criteria used in the selection and the evaluation of teachers (Doctoral dissertation, University of Iowa, 1977). Dissertation Abstracts International, 38, 1754-A.
- Al-Rubaiy, K. (1993). Five steps to better hiring. The Executive Educator, 15(8), 21-22.
- Arend, P. J. (1973). Teacher selection: The relationships between selected factors and the rated effectiveness of second-year teacher. Townson, MD: Baltimore County Board of Education. (ERIC Document Reproduction Service No. ED 087 102).
- Armstrong, C. (1988). Interview advice: Grill teachers about their track records. The Executive Educator, 10(9), 29.
- Ash, P. B. (1992). The 10-minute interview. The Executive Educator, 14(3), 40,42.
- Ayers, J. B., & Qualls, G. S. (1979). Concurrent and predictive validity of the National Teacher Examination. Journal of Educational Research, 73, 86-91.
- Bolton, D. L. (1973). Selection and evaluation of teachers. Berkeley, CA: McCutchan Publishing Corporation.
- Bolton, D. L. (1971). The use of simulation in educational administration. Columbus, OH: Charles E. Merrill Publishing Co.
- Boyles, N. L., & Engel, R. A. (1986). Finding teachers is tough; hiring the right ones is tougher. The Executive Educator, 8(8), 22-23.
- Borg, W. R., & Gall, M. D. (1989). Educational research: An introduction (5th ed.). White Plains, NY: Longman, Inc.
- Borich, G. D., & Fenton, K. S. (1977). The appraisal of teaching concepts and process. Reading, MS: Addison Wesley Publishing Company.
- Brandon, J. A. (1983). The teacher selection process as perceived by public school administrators in the state of Tennessee (Doctoral dissertation, The University of Tennessee, 1983). Dissertation Abstracts International, 43, 2839-A.

- Braun, J. A., Jr., Willems, A., Brown, M., & Green, K. (1987). A survey of hiring practices in selected school districts. Journal of Teacher Education, 38(2), 45-49.
- Bredeson, P. V. (1985). The teacher screening and selection process: A decision making model for school administrators. Journal of Research and Development in Education, 18(3), 8-15.
- Bredeson, P. V. (1982). The effects of letters of recommendation on the teacher selection interview process (Doctoral dissertation, The University of Wisconsin - Madison, 1982). Dissertation Abstracts International, 43, 984-A.
- Bridges, E. M. (1986). The incompetent teacher. Philadelphia, PA: The Falmer Press.
- Brooks, P. A. (1967). Teacher selection practices in Texas public schools (Doctoral dissertation, North Texas State University, 1967). Dissertation Abstracts International, 28, 3405-A.
- Brown, J. A. (1986). Preselection predictors of effective teaching performance (Omaha teacher interview, selection) (Doctoral dissertation, University of Florida, 1986). Dissertation Abstracts International, 47, 29-A.
- Browne, R. L. (1972). An analysis of the traditional teacher selection process and a classroom observation teacher selection process (Doctoral dissertation, Tempe Arizona State University, 1972). Dissertation Abstracts International, 32, 4268-A.
- Buffie, E. G. (1980). A job search model for the eighties. Bloomington, IN: Indiana University. (ERIC Document Reproduction Service No. ED 186 397).
- Burry, J. A., & Shaw, D. G. (1988). Teachers and administrators differ in assessing teacher effectiveness. Journal of Personnel Evaluation in Education, 2(1), 33-41.
- Caliendo, R. J. (1986). Selecting capable teachers. Spectrum, 4(1), 22-26.
- Callaghan, S. A. (1982). An evaluation of a data-based teacher selection process. (Doctoral dissertation, Northwestern University, 1982). Dissertation Abstracts International, 42, 3819-A.
- Cardozier, V. R. (1985). Recruiting and selecting college faculty. Austin, TX: Texas University. (ERIC Document Reproduction Service No. ED 263 831).



- Carrick, C. F., Jr. (1980). An analysis of teacher selection practices in local school systems of Georgia (Doctoral dissertation, University of Georgia, 1979). Dissertation Abstracts International, 40, 4317-A.
- Cherrington, D. J. (1987). Personnel management (2nd ed.). Dubuque, IA: Wm. C. Brown Publishers.
- Clark, S. (1992). No way but up. The Executive Educator, 14(12), 24-26.
- Cockrill, J. E. (1983). Procedures and criteria used by Texas public school administrative personnel specialists in the selection of teachers (Doctoral dissertation, East Texas State University, 1982). Dissertation Abstracts International, 43, 3467-A.
- Cook, R. E. (1982). Chippewa valley local validation study of the teacher perceiver interview process (Doctoral dissertation, Wayne State University, 1982). Dissertation Abstracts International, 42, 4662-A.
- Cox, D. L. (1981). The board's responsibility for attracting and landing the best teaching staff. (ERIC Document Reproduction Service No. ED 207 142).
- Dale, J. (1991). Leave hiring to the experts. The Executive Educator, 13(4), 20-21.
- Davey, B. (1991). Evaluating teacher competence through the use of performance assessment tasks: an overview. Journal of Personnel Evaluation in Education, 5(2), 121-132.
- Diekmann, R. G. (1981). A study of the relationships between predicted effectiveness of teacher candidates and actual effectiveness after one year of teaching experience (Doctoral dissertation, University of Nebraska, 1981). Dissertation Abstracts International, 42, 3823.
- Evertson, C. M., & Burry, J. A. (1989). Capturing classroom content: The observation system as lens for assessment. Journal of Personnel Evaluation in Education, 2(4), 297-320.
- Ferguson, G. A., & Takane, Y. (1989). Statistical analysis in psychology and education (6th ed.). New York: McGraw-Hill Book Company.
- Fiske, E. B. (1978). Teachers are asked for proof of skills. The New York Times, 128(44,036), C1, C5.
- Fitzpatrick, R., & Morrison, E. J. (1971). Performance and product evaluation. In R. L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, DC: American Council on Education.

- Fuhr, D. L. (1977). Selecting first year teachers. NASSP Bulletin, 61(406), 57-59.
- Galbo, J. J., & Diekman, J. E. (1985). The criteria utilized in the selection of beginning teachers: An extension and partial replication. (ERIC Document Reproduction Service No. ED 275 080).
- Gerber, M. (1968). The development and evaluation of a weighted application blank for use in the selection of elementary school teachers (Doctoral dissertation, University of California-Berkeley, 1968). Dissertation Abstracts International, 28, 4409-4410.
- Gerwin, D. (1974). The employment of teachers. Berkeley, CA: McCutchan Publishing Corporation.
- Ghysels, M. (1988). A study of the perspectives, assumptions, and beliefs of school administrators towards teacher lesson demonstrations within the context of the teacher selection process. Unpublished master's thesis. California State University at Hayward, Hayward, CA.
- Gilbert, H. B., & Lang G. (1967). Teacher selection methods. New York: New York City Board of Education. (ERIC Document Reproduction Service No. ED 014 469).
- Ginott, H. G. (1972). Teacher and child. New York: The Macmillan Company.
- Grover, B. W. (1991). The teacher assessment dilemma: What is versus what ought to be! Journal of Personnel Evaluation in Education, 5(2), 103-119.
- Haefele, D. L. (1992). Evaluating teachers: An alternative model. Journal of Personnel Evaluation in Education, 5(4), 335-345.
- Haertel, E. H. (1986). The valid use of student performance measures for teacher evaluation. Educational Evaluation and Policy Analysis, 8(1), 45-60 (ERIC Document Reproduction Service No. EJ 350 184).
- Harrington-Lueker, D. (1992). Up from the ashes with project Phoenix. The Executive Educator, 14(12), 26.
- Herman, J. J., & Stephens, G. M. (1987). Follow this eight-step plan for hiring first-rate teachers. The Executive Educator, 9(5), 24-25.
- Heynderickz, J. J. (1987). Teacher selection. Research Roundup, 3(3), 6-11.

- Hill, W. D. (1978). Reference checking--a comparison of two methods (Doctoral dissertation, University of Texas at Austin, 1978). Dissertation Abstracts International, 39, 1955-A.
- Ishee, W. W. (1982). Analysis of teacher selection procedures in suburban school districts in Harris county, Texas (Doctoral dissertation, Texas A&M University, 1982). Dissertation Abstracts International, 42, 2950-A.
- Jackson, J. A. (1983). Criteria for selecting public school teachers in the state of Oklahoma (Doctoral dissertation, East Texas State University, 1983). Dissertation Abstracts International, 44, 1644-A.
- Jacobson, L. S., & Pecheone, R. L. (1991). Connecticut teacher assessment center (CONNTAC) program: Assessing professional knowledge of beginning teachers. Journal of Personnel Evaluation in Education, 5(2), 205-226.
- Jensen, M. C. (1987). How to recruit, select, induct, and retain the very best teachers. School Management Digest Series, 32 (ERIC Document Reproduction Service No. 279 056).
- Jensen, M. C. (1986). Recruiting and selecting the most capable teachers. OSSC Bulletin, 29(9), 1-28 (ERIC Document Reproduction Service No. ED 269 885).
- Johannsen, K. L. (1982). The relationship of a unique teacher selection method to student achievement (Doctoral dissertation, Peabody College for Teachers of Vanderbilt University, 1982). Dissertation Abstracts International, 42, 4671-A.
- Kelly, J. A. (1989). Examination of SERs as used in Iowa school districts for evaluating teachers (Doctoral dissertation, Drake University, 1989). Dissertation Abstracts International, 50, 1873-A.
- Kirkland, S. E. (1979). The development and implementation of a system to facilitate the interviewing of potential teachers by heads of elementary administrative units in the Sacramento City Unified School District. (ERIC Document Reproduction Service No. ED 196 821).
- Kopetz, B. E. (1982). The effect of selected characteristics of first-time applicants for instrumental music positions on teacher employment decisions (Doctoral dissertation, Indiana University, 1982). Dissertation Abstracts International, 42, 3490-A.

- Leshner, M. R., & Wade, S. (1972). A study of teacher employment practices in Iowa, Kansas, Missouri, and Nebraska, Maryville, MO: Northwest Missouri State University. (ERIC Document Reproduction Service No. 063 264).
- LeTendre, B. G. (1989). The use of work samples in hiring teachers (Doctoral dissertation, Stanford University, 1989) Dissertation Abstracts International, 50, 2006-A (University Microfilms No. 8919398).
- Loehr, P. (1986). Relationships between teacher interview scores and on-the-job performance. (ERIC Document Reproduction Service No. ED 277 118).
- Luthy, T. F. (1984). Policies and procedures for selection of Missouri school teachers (Doctoral dissertation, University of Missouri - Columbia, 1984) Dissertation Abstracts International, 44, 1995-A (University Microfilms No. 8326825).
- Manatt, R. P. (1982). Teacher performance evaluation - practical application of research. Ames, Iowa: ISU, Research Institute for Studies in Education (ERIC Document Reproduction Service No. ED 225 275).
- McGarity, J. L. (1987). An evaluation of selected criteria used for the selection of qualified teacher applicants in the Guinnett county (Georgia) public school district (Doctoral dissertation, Georgia State University, 1987). Dissertation Abstracts International, 48, 272-A.
- McGreal, T. L. (1990). The use of rating scales in teacher evaluation: Concerns and recommendations. Journal of Personnel Evaluation in Education, 4(1), 41-58.
- McIntyre, K. E. (1974). The way it was/is. The National Elementary Principal, 53, 30-34.
- Merritt, D. (1971). Attitude congruency and selection of teacher candidates. Administrator's Notebook, 19(6), 1-4.
- Mickler, M. L., & Solomon, G. L. (1986). Beyond credentials in teacher selection: A second look at the Omaha teacher interview. North Central Association Quarterly, 60(3), 398-402.
- Millman, J., & Darling-Hammond, L. (1990). The new handbook of teacher evaluation: Assessing elementary and secondary school teachers. Newbury Park, CA: Sage Publications, Inc.
- Moore, A. J. (1987). Selecting outstanding faculty: Indicators of quality. Journal of Teacher Education, 38(1), 43-47.

- Mortaloni, R. (1974). School administrators evaluate the letter of reference and selected recruiting practices. (ERIC Document Reproduction Service No. ED 099 965).
- Murphy, J. (1987). Teacher evaluation: A comprehensive framework for supervisors. Journal of Personnel Evaluation in Education, 1(2), 157-180.
- Nalley, B. J. (1971). A descriptive survey of the recruitment and selection process of teachers for the District of Columbia and a comparison of procedures used in selected school systems of comparable size (Doctoral dissertation, George Washington University, 1971). Dissertation Abstracts International, 32, 3626A.
- Nicholson, E. W., & McInerney, W. D. (1988). Hiring the right teacher: A method for selection. NASSP Bulletin, 72(511), 88-92.
- Orebaugh, F. E. (1974). The analysis of selected employment criteria and their relationship to the performance ratings of teachers (Doctoral dissertation, University of Cincinnati, 1974). Dissertation Abstracts International, 35, 2597A.
- Overman, W. F. (1982). The importance of the teacher perceiver interview as a teacher selection criterion to administrators in relationship to their leadership style (Doctoral dissertation, University of Northern Colorado, 1982). Dissertation Abstracts International, 42, 3373-A.
- Place, A. W., & Kowalski, T. J. (1993). Principal ratings of criteria associated with teacher selection. Journal of Personnel Evaluation in Education, 7(4), 291-300.
- Pounder, D. G. (1989). Improving the predictive validity of teacher selection decisions: Lessons from teacher appraisal. Journal of Personnel Evaluation in Education, 2(2), 141-150.
- Raths, J. (1989). Reformers' visions of tomorrow's teachers: A U.S.A. perspective. Childhood Education, 65, 263-267.
- Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology, 35, 1-62.
- Ross, J. R. (1991). Give us the power. The Executive Educator, 13(4), 18-19.
- Schmitt, N., Gooding, R., Noe, R. A., & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

- Scriven, M. (1990). Can research-based teacher evaluation be saved? Journal of Personnel Evaluation in Education, 4(1), 19-32.
- Seiferth, B., & Purcell, T. D. (1979). Administrators' criteria for hiring practices. (ERIC Document Reproduction Service No. ED 177 096).
- Shanoski, L. A., & Hranitz, J. R. (1989). An analysis of characteristics of outstanding teachers and the criteria used by colleges and universities to select future teachers. (ERIC Document Reproduction Service No. ED 312 236).
- Shelton, M. M. (1989). Teacher selection: A training model for principals. (ERIC Document Reproduction Service No. ED 311 533).
- Skillett, J. D., & Tompkins, L. D. (1984). The hiring process as a quality control measure. Myth or reality? (ERIC Document Reproduction Service No. ED 262 012).
- Smedberg, C. A. (1984). An examination of the validity of the empathy interview as an instrument for teacher selection (Florida; beginning teacher) (Doctoral dissertation, University of South Florida, 1984). Dissertation Abstracts International, 45, 1263-A.
- Smith, J. D. (1982). A study of correlations for pretest and posttest scores on a structured interview for teacher selection, and student and administrator ratings of teachers (Doctoral dissertation, The University of Nebraska - Lincoln, 1982). Dissertation Abstracts International, 42, 4681-A.
- Smith, J. M. (1980). Teaching selection procedures. The Clearing House, 53(7), 312-314.
- Solomon, G. L. (1982). The Omaha teacher pre-employment interview as a predictor of job performance of selected new teachers in the Jefferson county school system (Doctoral dissertation, The University of Alabama, 1982). Dissertation Abstracts International, 42, 3833-A.
- Stone, B. (1972). The effect of classroom observations on teacher selection decisions (Doctoral dissertation, University of Oklahoma, 1972). Dissertation Abstracts International, 24, 939-A.
- Teitelbaum, D., & McLaughlin, J. H. (1979). National survey of teacher and supervisor selection procedures in fifty-four large public school systems in the United States. New York, NY: New York City board of education. (ERIC Document Reproduction Service No. ED 182 271).

- Thacker, J. A. (1974). A study of the relationship between principal' estimates of teaching efficiency and scores on the national teacher exam, academic averages and supervisors' estimates of potential for selected teachers in North Carolina (Doctoral dissertation, University of North Carolina at Chapel Hill, 1974). Dissertation Abstracts International, 26, 1462A.
- Theodossin, E. (1983). Selection through interviewing: Entrance procedures in teacher recruitment. Blagdon, England: Coombe Lodge Research Report. (ERIC Document Reproduction Service No. ED 267 036).
- Title, D. G. (1995). How to hire a teacher. Sacramento, CA: AASPA
- Tomala, G., & Weinland, T. (1991). Social studies: Connecticut's teacher assessment center exercises. Journal of Personnel Evaluation in Education, 5(2), 191-204.
- Tractenberg, P. L. (1973). Testing the teacher: How urban school districts select their teachers and supervisors. New York: Agathon Press, Inc.
- Travers, P. D. (1989). Preparing for teacher employment: An analysis of job applications. The Clearing House, 62(6), 263-265.
- Underwood, K. E. (1961). An analysis of letters of recommendation (Doctoral dissertation, Indiana University, 1961). Dissertation Abstracts International, 22, 3079.
- Vanderheiden, D. L. (1982). A study of the processes utilized in recruitment and selection of teachers in selected school districts: A planning model (Doctoral dissertation, The University of Nebraska - Lincoln, 1982). Dissertation Abstracts International, 42, 4683-A.
- Vann, A. S. (1991). At least listen to the principal. The Executive Educator, 13(4), 21-23.
- Vann, A. S. (1989, September). A principal's guide to the hiring and induction of new teachers. Principal, 26-29.
- Vavrus, L., & Collins A. (1991). Portfolio documentation and assessment center exercises: A marriage made for teacher assessment. Teacher Education Quarterly, 18(3), 13-29.
- Voss, G. C. (1983). The effect of age, position and amount of resume information on the selection of teacher candidates (Doctoral dissertation, The University of Wisconsin - Madison, 1983). Dissertation Abstracts International, 44, 1288-A.

- Wasicsko, M. M. (1977). A research-based teacher selection instrument. (ERIC Document Reproduction Service No. ED 193 193).
- Webster, W. J. (1988). Selecting effective teachers. Journal of Educational Research, 81(4), 245-253.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70(9), 703-713.
- Wise, A. E., Darling-Hammond, L., & Berry, B. (1988). Selecting teachers: The best, the known, and the persistent. Educational Leadership, 45(5), 82-85.
- Wise, A. E., Darling-Hammond, L., & Berry, B. (1987). Effective teacher selection: From recruitment to retention. Research report. Santa Monica, CA: Rand Corporation.
- Young, I. P., & McMurry, B. R. (1986). Effects of chronological age, focal position, quality of information and quantity of information on screening decisions for teacher candidates. Journal of Research and Development in Education, 19(4), 1-9.
- Young, J., & Elliott, M. (1986). Teacher screening and selection processes: A survey of nine Winnipeg school divisions. (ERIC Document Reproduction Service No. ED 277 093).



## Appendix A

School District A  
SUMMATIVE EVALUATION REPORT

Teacher \_\_\_\_\_  
 School \_\_\_\_\_  
 Assignment \_\_\_\_\_  
 Date of Summative Evaluation Conference \_\_\_\_\_

"We recognize that a quality staff is the key to providing quality education. We are dedicated to the development of the professional skills and human potential of all staff members." -- from School District A Mission Statement.

"Instructional improvement is a priority of School District A. Improvement based on performance evaluation encourages high student achievement as well as the growth of pride and excellence in the instructional program." -- from School District A Philosophy of Performance Evaluation.

Definitions for Levels of Performance on the Summative Evaluation Report

"Performance meets district standards" - There is evidence that both the quality and the consistency of performance regularly meets or occasionally exceeds the expectations of the districts.

"Performance exceeds district standards" - There is evidence that both the quality and the consistency of performance regularly exceeds the standard of the district. Performance in this area is a model for others. Detailed evidence of this commendation is attached to the Summative Evaluation Report.

"Performance does not meet district standards" - There is evidence that both the quality and the consistency of performance does not meet the expectations of the district in this area. Detailed evidence of this commendation is attached to the Summative Evaluation Report.

Teacher's Signature \_\_\_\_\_ Date \_\_\_\_\_

(The teacher's signature only indicates that the evaluation report has been reviewed with him or her. It does not necessarily mean that he or she agrees with the evaluation.)

Principal's Signature \_\_\_\_\_ Date \_\_\_\_\_

Teacher's Signature \_\_\_\_\_ Date \_\_\_\_\_

<b>CRITERIA</b>	<b><u>RATINGS</u></b>	<b><u>RATINGS</u></b>
<b>Performance Area I: PRODUCTIVE TEACHING TECHNIQUES</b>		
A. Develops lesson plans which are congruent with district approved course outlines or guides.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.
B. Implements lesson plan.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.
C. Uses appropriate techniques to motivate students.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.
D. Communicates effectively with students.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.
E. Evaluates pupil progress.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.
F. Demonstrates an appropriate knowledge of curriculum and subject matter.	<input type="checkbox"/> Performance meets district standards. <hr/> <hr/> <hr/>	<input type="checkbox"/> Performance exceeds the district standards.  <input type="checkbox"/> Performance does <u>not</u> meet the district standards.

Teacher's Signature \_\_\_\_\_ Date \_\_\_\_\_

**CRITERIA****RATINGS****RATINGS****Performance Area I:**

- G. Sets high expectations for student achievement.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

**Performance Area II:****CLASS MANAGEMENT**

- A. Organizes the educational setting.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- B. Demonstrates evidence of personal organization.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- C. Sets high standards for student behavior.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- D. Organizes students for effective instruction.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

**Performance Area III:****EFFECTIVE INTERPERSONAL RELATIONS**

- A. Demonstrates effective interpersonal relationships.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

Teacher's Signature \_\_\_\_\_ Date \_\_\_\_\_

**CRITERIA****RATINGS****RATINGS**Performance Area III:**EFFECTIVE INTERPERSONAL RELATIONS**

- B. Demonstrates a sensitive awareness of the needs of students.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- C. Demonstrates a sensitive awareness of the needs of colleagues.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- D. Demonstrates a sensitive awareness of the needs of parents and other adults.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- E. Promotes positive self-concept.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

Performance Area IV:**PROFESSIONAL RESPONSIBILITIES**

- A. Adheres to school regulations and policies.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

- B. Establishes goals and initiates activities aimed at professional growth and development.

☐ Performance meets district standards.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ Performance exceeds the district standards.

☐ Performance does not meet the district standards.

Teacher's Signature \_\_\_\_\_ Date \_\_\_\_\_